An Update on Proxy System Modeling and Model-Data Comparison

progress, challenges, and applications

Sylvia Genevieve Dee

Brown University, Department of Earth, Environmental, and Planetary Sciences
Voss Postdoctoral Fellow, Institute at Brown for Environment and Society

PAGES2k-PMIP3 Hydroclimate Workshop
// LDEO, NY // June 3rd , 2016

What is a Proxy System Model?

Proxy system models are pretty cool.

A brief outline:
developing best practices for model-data comparison

- **PRoxY System Modeling Progress:**
  - Open-source, public PSM tools: PRYSM v1.0 and 2.0
- **Advanced Applications in Data-Model Comparison:**
  - Data Assimilation and Paleoclimate Reanalyses with PSMs
  - Investigating parametric uncertainties ~ checking our understanding of the proxy system
  - Data-model comparison in the frequency domain using PSMs

Building a PSM code package

- Many PSMs are completed or in development
- We need a common model development framework
  - Encourage feedback, expand use
  - Format/structure for building and comparing PSMs
- **Goals:**
  - Adaptable, modular, and public
  - Consolidate inter-model redundancies
  - Open Source

python

ObsPy
A Python Framework for Seismology

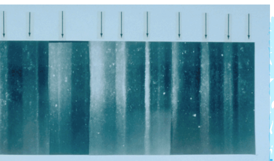
astropy
A Community Python Library for Astronomy

Biopython Project
http://www.biopython.org

Nitime: time-series analysis for neuroscience

PRYSM:

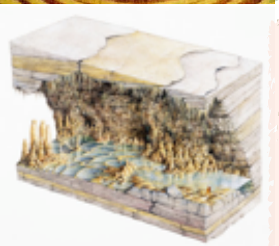
a public PSM development platform



ICE CORES

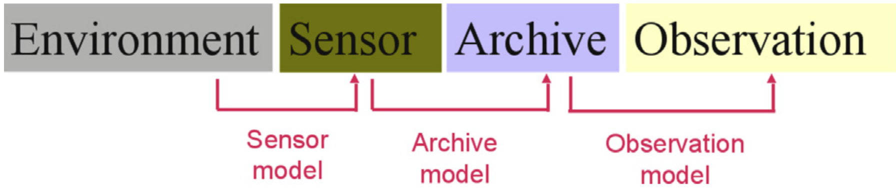
CORALS

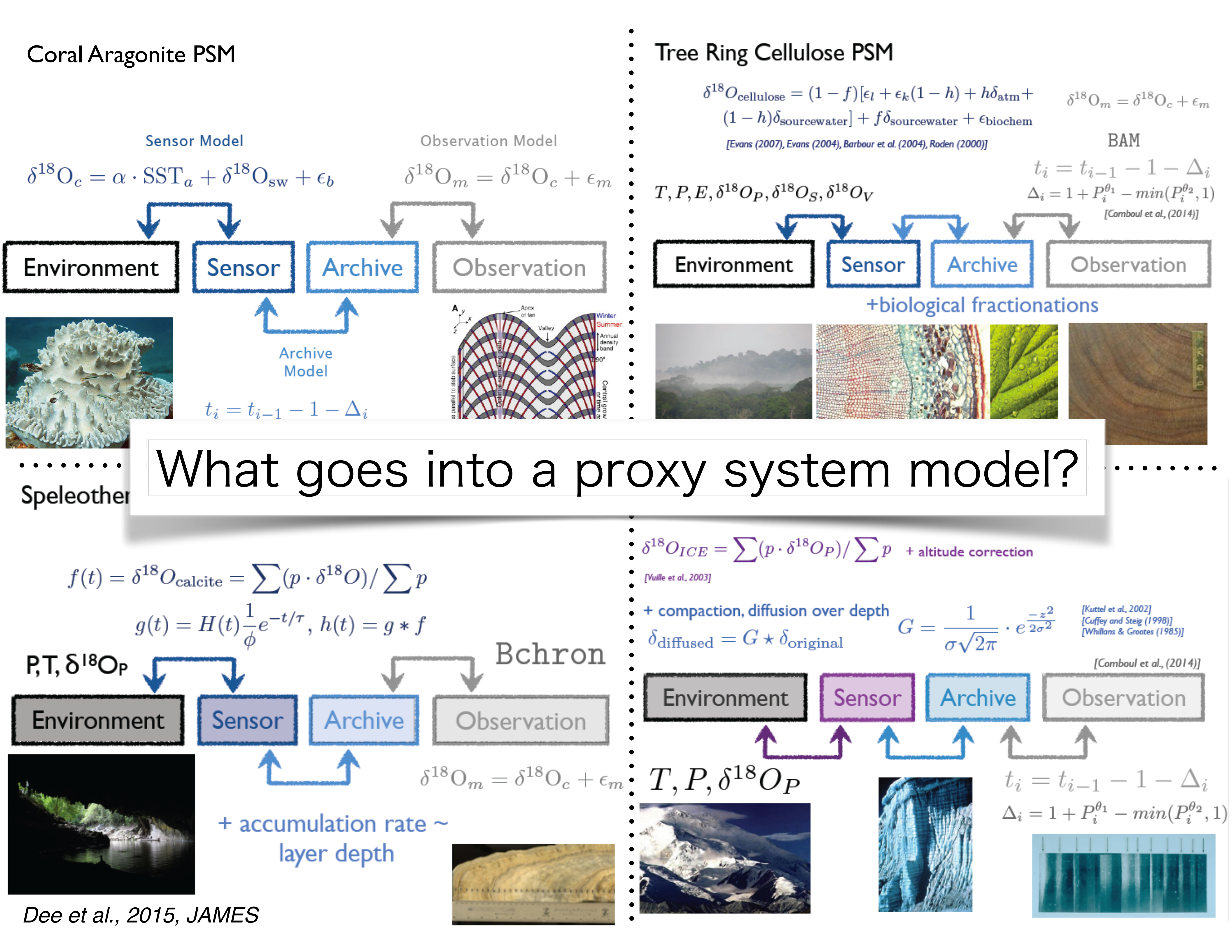
TREE RING CELLULOSE



SPELEOTHEMS

- synthesized from previously published work
- current availability: high-resolution water isotope systems
    - enabled by recent water isotope-enabled modeling efforts, SWING2 -> full forward modeling scheme
- Sub-model framework, as in [Evans et al. QSR, 2013]





What goes into a proxy system model?
Coral Aragonite PSM
Sensor Model
Observation Model
$\delta^{18}O_c = \alpha \cdot SST_a + \delta^{18}O_{sw} + \epsilon_b$
$\delta^{18}O_m = \delta^{18}O_c + \epsilon_m$
Environment
Sensor
Archive
Observation
Archive Model
$t_i = t_{i-1} - 1 - \Delta_i$
Tree Ring Cellulose PSM
$\delta^{18}O_{cellulose} = (1-f)[\epsilon_l + \epsilon_k(1-h) + h\delta_{atm} + (1-h)\delta_{sourcewater}] + f\delta_{sourcewater} + \epsilon_{biochem}$
[Evans (2007), Evans (2004), Barbour et al. (2004), Roden (2000)]
$\delta^{18}O_m = \delta^{18}O_c + \epsilon_m$
BAM
$t_i = t_{i-1} - 1 - \Delta_i$
$\Delta_i = 1 + P_i^{\theta_1} - min(P_i^{\theta_2}, 1)$
[Comboul et al., (2014)]
$T, P, E, \delta^{18}O_P, \delta^{18}O_S, \delta^{18}O_V$
Environment
Sensor
Archive
Observation
+biological fractionations
Speleother
$f(t) = \delta^{18}O_{calcite} = \sum(p \cdot \delta^{18}O) / \sum p$
$g(t) = H(t)\frac{1}{\phi}e^{-t/\tau},\ h(t) = g * f$
Bchron
P,T, $\delta^{18}O_P$
Environment
Sensor
Archive
Observation
$\delta^{18}O_m = \delta^{18}O_c + \epsilon_m$
+ accumulation rate ~ layer depth
Dee et al., 2015, JAMES
$\delta^{18}O_{ICE} = \sum(p \cdot \delta^{18}O_P) / \sum p$
+ altitude correction
[Vuille et al., 2003]
+ compaction, diffusion over depth
$\delta_{diffused} = G \star \delta_{original}$
$G = \frac{1}{\sigma\sqrt{2\pi}} \cdot e^{\frac{-z^2}{2\sigma^2}}$
[Kuttel et al., 2002]
[Cuffey and Steig (1998)]
[Whillans & Grootes (1985)]
[Comboul et al., (2014)]
Environment
Sensor
Archive
Observation
$T, P, \delta^{18}O_P$
$t_i = t_{i-1} - 1 - \Delta_i$
$\Delta_i = 1 + P_i^{\theta_1} - min(P_i^{\theta_2}, 1)$

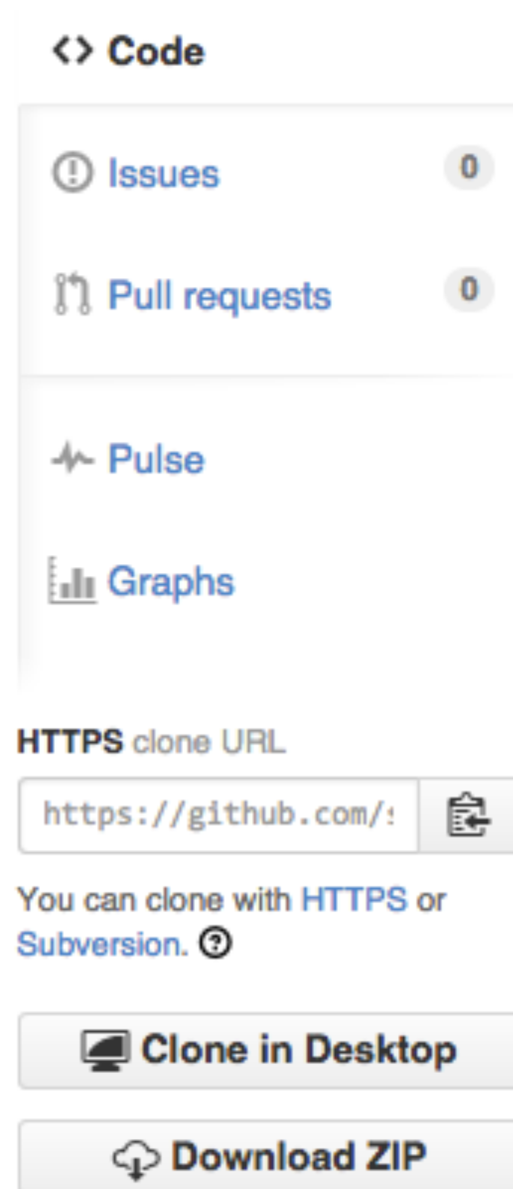
PRYSM on GitHub & Installation

`https://github.com/sylvia-dee/PRYSM`

git clone https://github.com/sylvia-dee/PRYSM.git

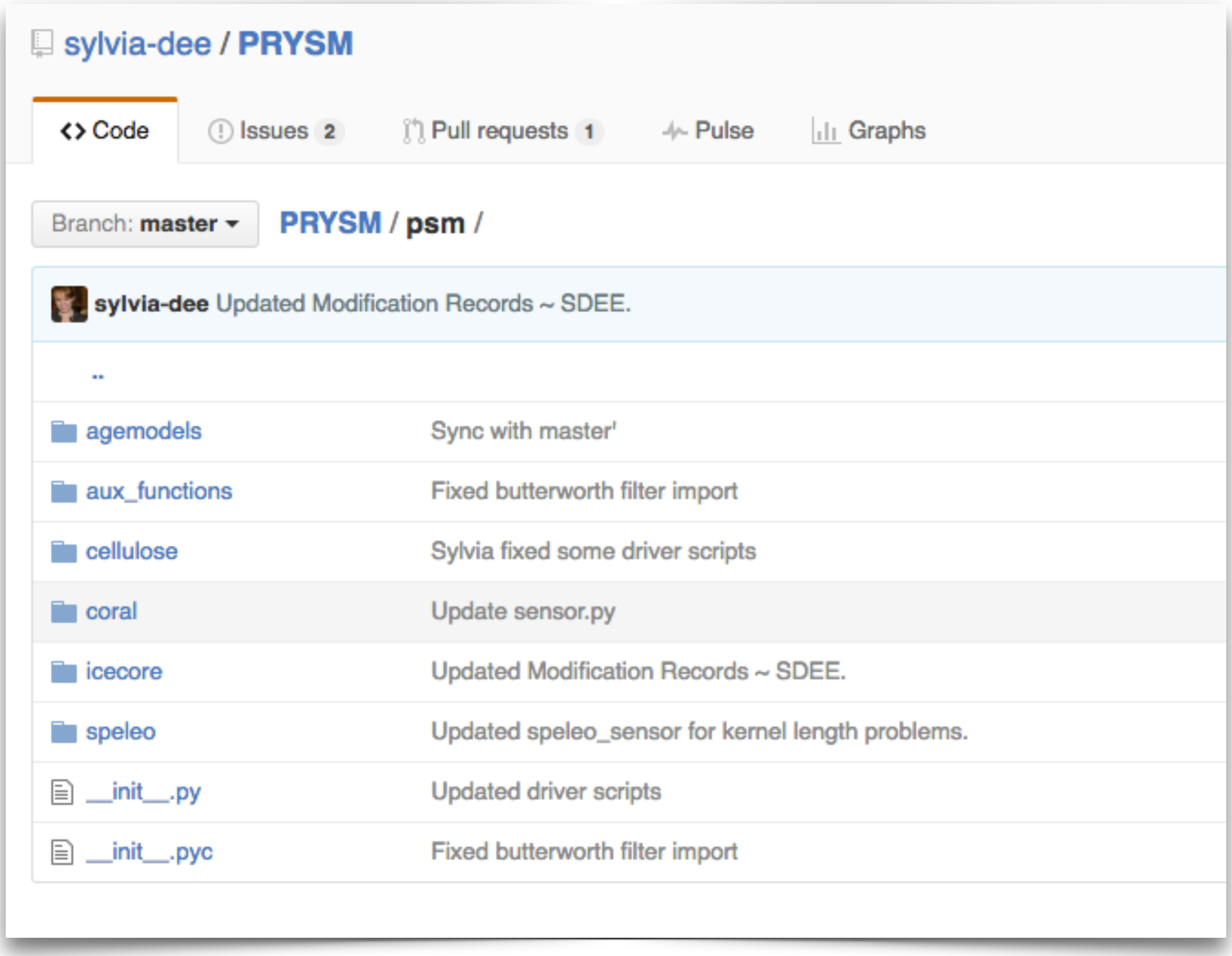
1. Download zipfile
2. In working directory: use Python

```
>> python setup.py install
>> import psm
```

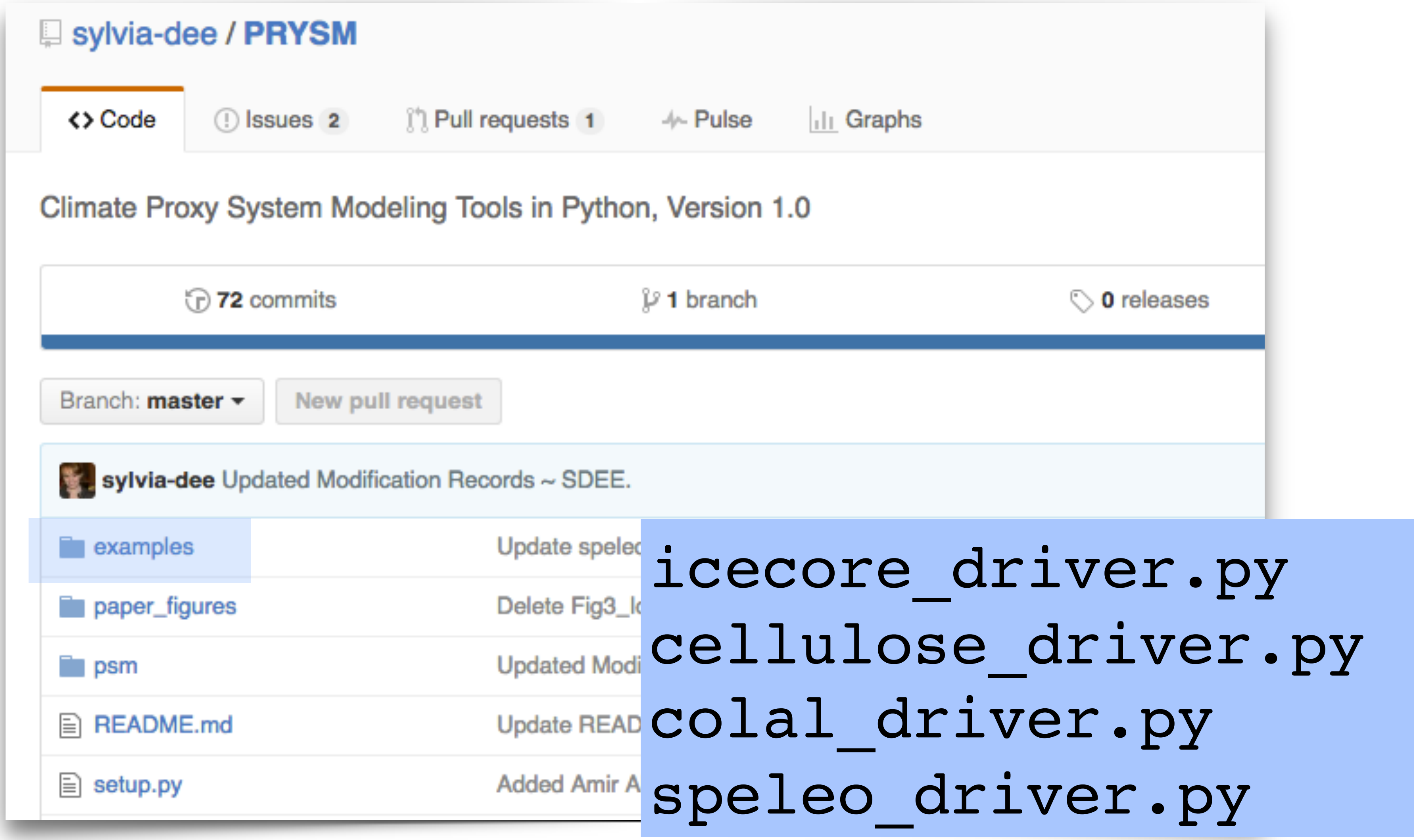


Dee et al., 2015, JAMES

PRYSM



sylvia-dee / PRYSM
Code
Issues 2
Pull requests 1
Pulse
Graphs
Branch: master
PRYSM / psm /
sylvia-dee Updated Modification Records ~ SDEE.
..
agemodels
Sync with master'
aux_functions
Fixed butterworth filter import
cellulose
Sylvia fixed some driver scripts
coral
Update sensor.py
icecore
Updated Modification Records ~ SDEE.
speleo
Updated speleo_sensor for kernel length problems.
__init__.py
Updated driver scripts
__init__.pyc
Fixed butterworth filter import



>> these driver scripts walk you through running the PSM sub-models in succession.

you can run files in iPython to execute step-by-step...

```
#==========================================================================
# E3. CALL SENSOR MODEL
#==========================================================================
print 'Running sensor model...'
# 4. Apply icecore_sensor to extract precipitation-weighted d18o record for each core
#    and compute altitude, temperature corrections. (Please see docstring icecore_sensor).

d18O      = deltaP       # your dataset loaded here
alt_diff = 3524.0        # alt_diff at location (m) (THIS IS FOR SPEEDY-QUELCCAYA)
d18Oice   = icecore_sensor(time,d18O,alt_diff)

# returns: icecore

#==========================================================================
# E4. CALL ARCHIVE MODEL
#==========================================================================
print 'Running archive model...'
# This archive model will calculate diffusion and compaction
# (Please see docstrings: diffusivity, icecore_diffuse)

# NOTE: tester file has accumulation in meters per year. Below is optional unit conversion.

#accum=accum*365.0       # multiple by 365 days to get yearly accumulation in mm
#b = accum/1000.         # convert mm/yr to m/yr, accumulation rate (e.g. 1.3 m/year)

b=accum
core_length=np.cumsum(b)
depth = core_length[-1]
```

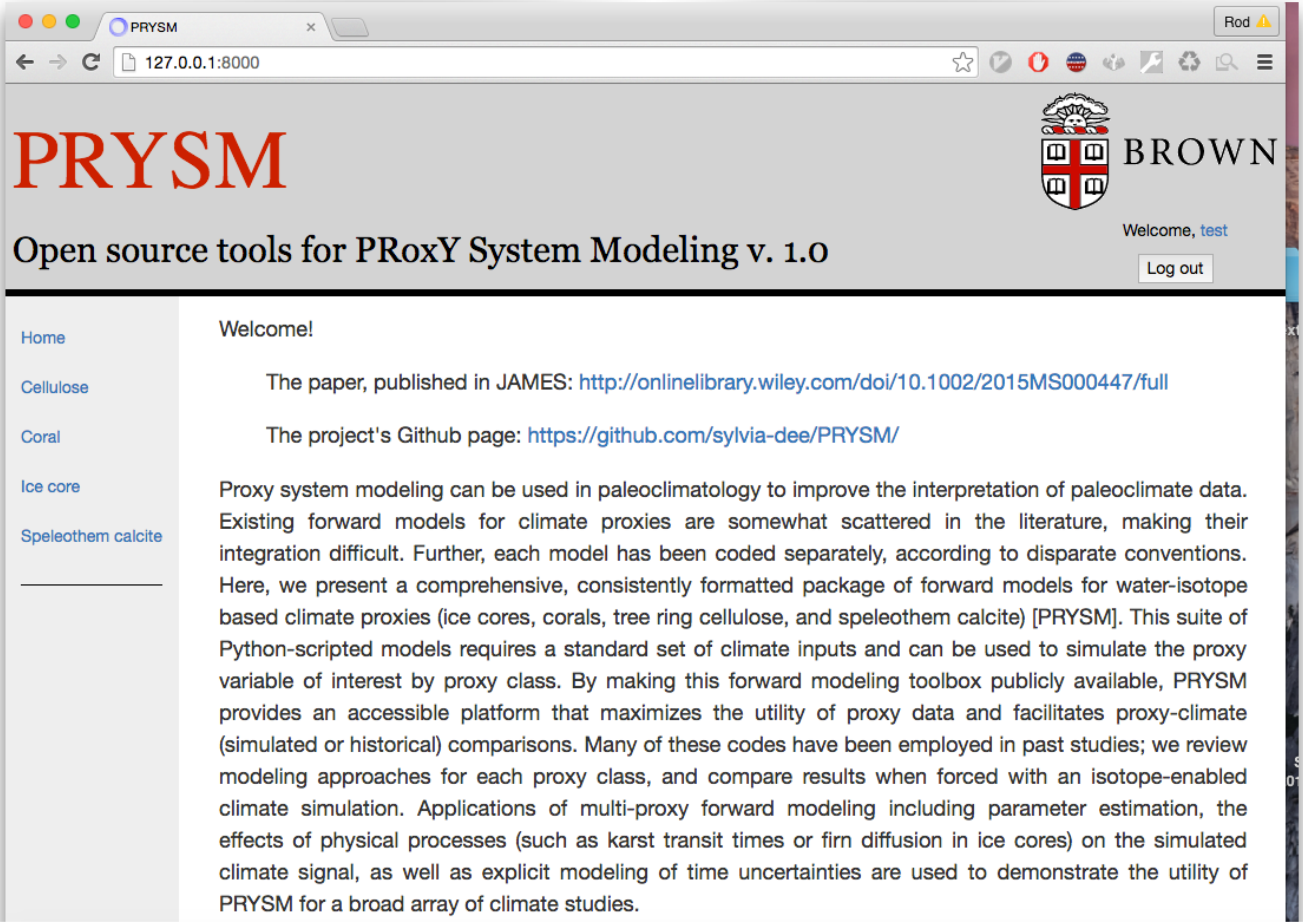
```python
def diffusivity(rho,T=250,P=0.7,rho_d=822,b=0.25):

    '''
    DOCSTRING: Function 'diffusivity'
    Description: Calculates diffusivity (in m^2/s) as a function of density.

    Inputs:
    P: Ambient Pressure in Atm
    T: Temperature in K
    rho: density profile (kg/m^3)
    rho_d: 822 kg/m^2 [default], density at which ice becomes impermeable to diffusion

    Defaults are available for all but rho, so only one argument need be entered.

    Note values for diffusivity in air:

    D16 = 2.1e-5*(T/273.15)^1.94*1/P
    D18 = D16/1.0285
    D2 = D16/1.0251
    D17 = D16/((D16/D18)^0.518)

    Reference: Johnsen et al. (2000): Diffusion of Stable isotopes in polar firn and ice:
    the isotope effect in firn diffusion

    '''
    import numpy as np
    import scipy
    from scipy import integrate
    import matplotlib.pyplot as plt

```

PRYSM

Perhaps Python is not for you..


KEEP
CALM.
THERE'S
AN APP
FOR THAT

PRYSM GUI



PRYSM

127.0.0.1:8000

Rod

PRYSM

BROWN

Open source tools for PRoxY System Modeling v. 1.0

Welcome, test

Log out

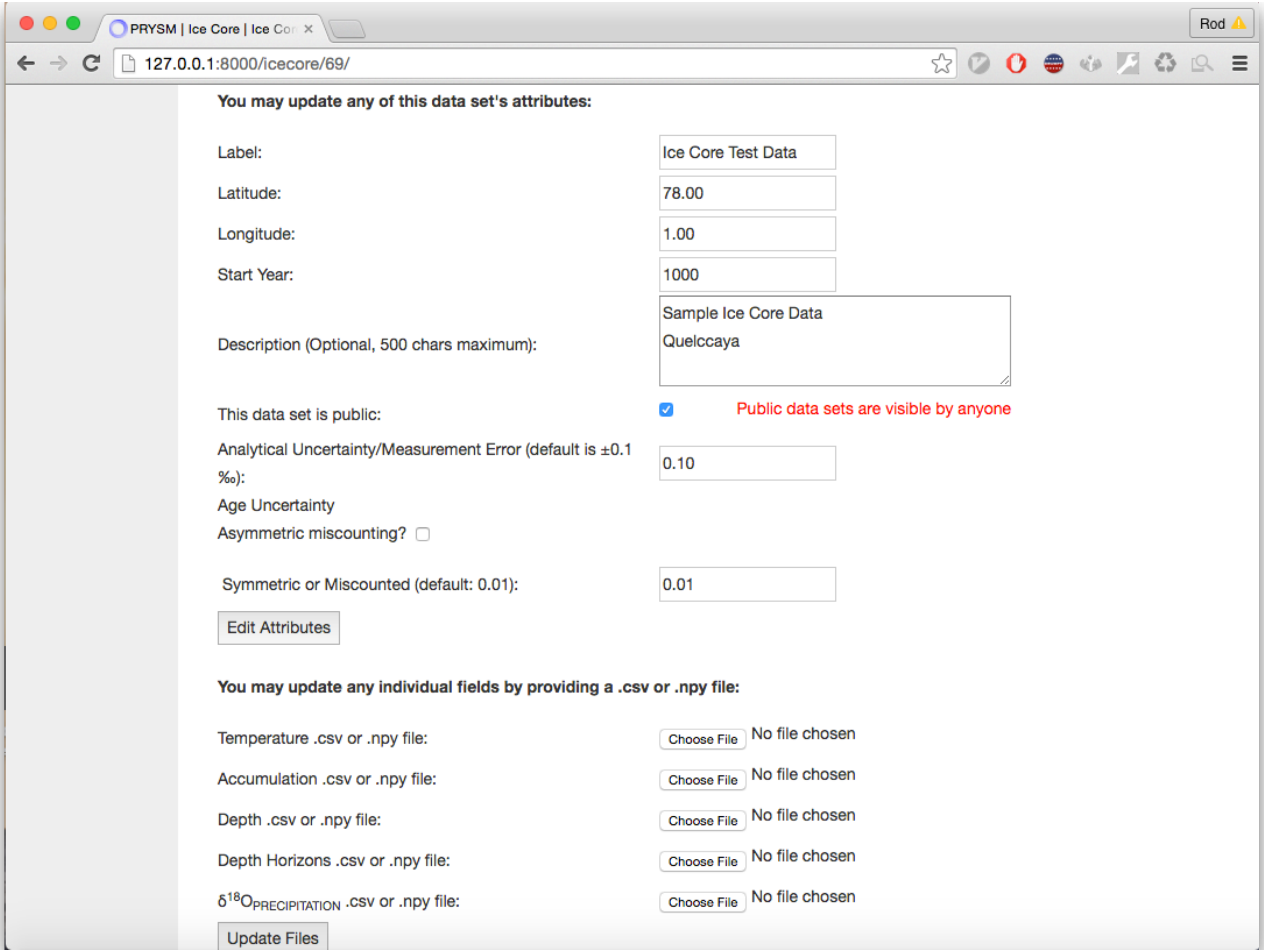
- Home
- Cellulose
- Coral
- Ice core
- Speleothem calcite

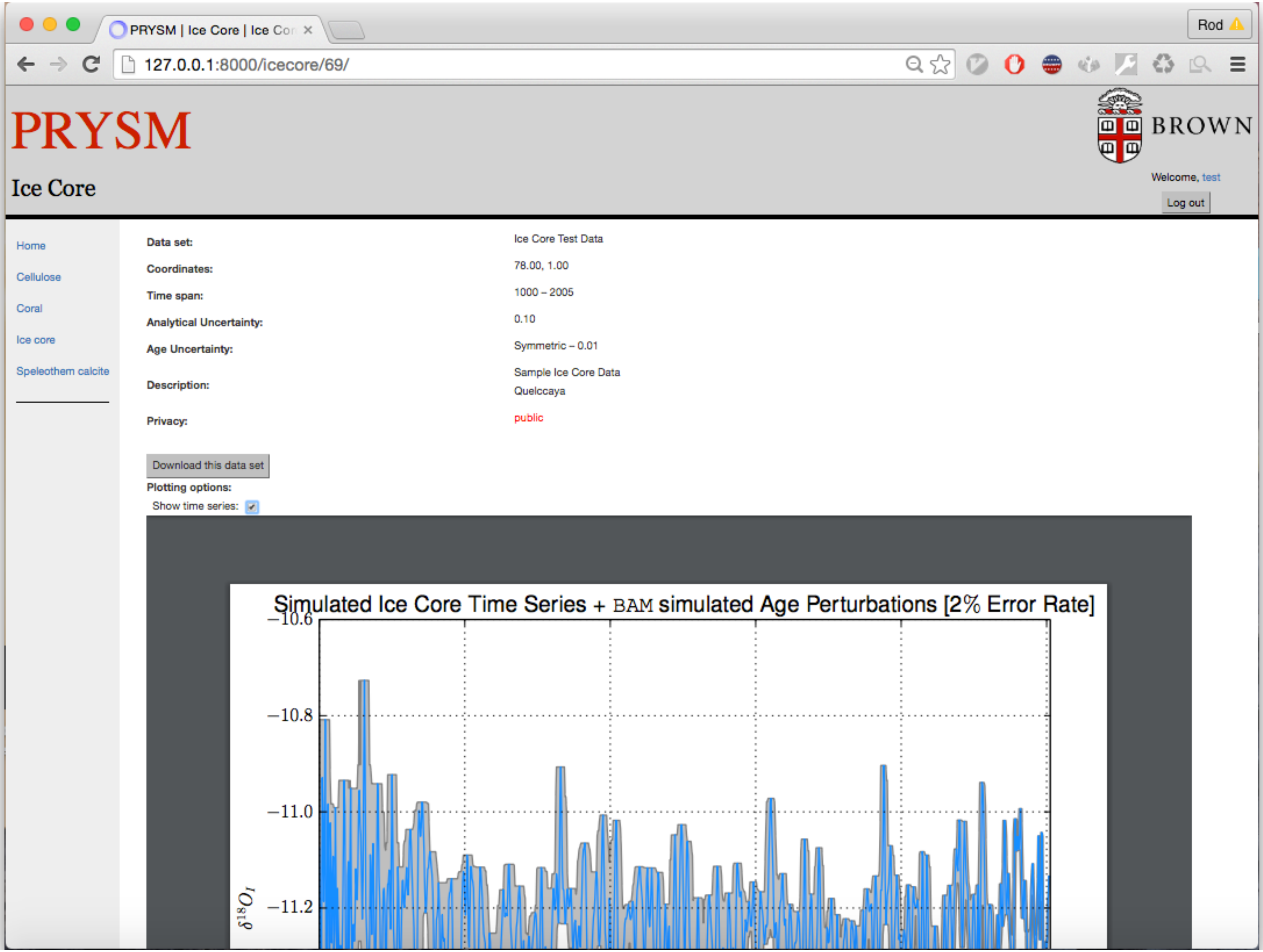
Welcome!

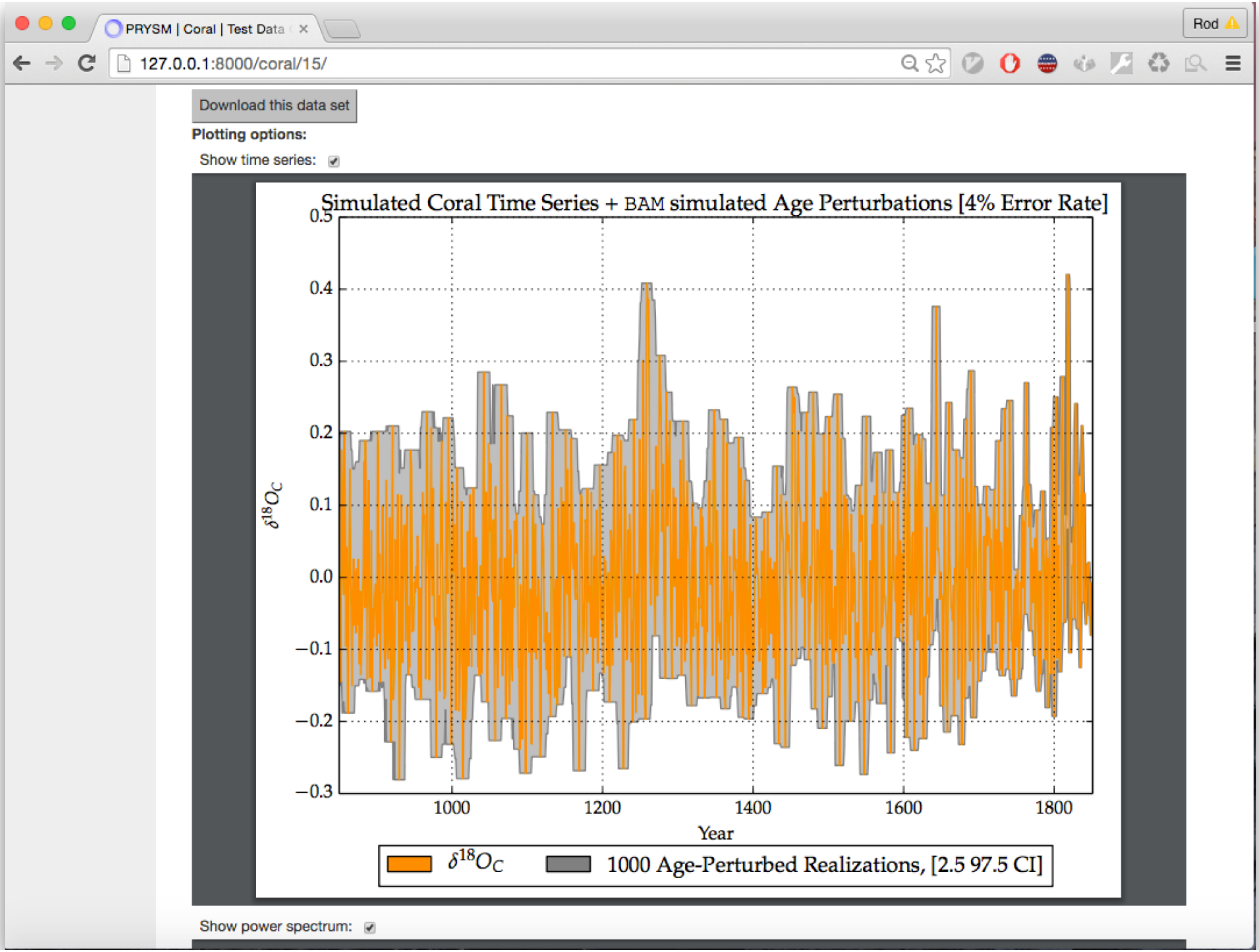
The paper, published in JAMES: http://onlinelibrary.wiley.com/doi/10.1002/2015MS000447/full

The project's Github page: https://github.com/sylvia-dee/PRYSM/

Proxy system modeling can be used in paleoclimatology to improve the interpretation of paleoclimate data. Existing forward models for climate proxies are somewhat scattered in the literature, making their integration difficult. Further, each model has been coded separately, according to disparate conventions. Here, we present a comprehensive, consistently formatted package of forward models for water-isotope based climate proxies (ice cores, corals, tree ring cellulose, and speleothem calcite) [PRYSM]. This suite of Python-scripted models requires a standard set of climate inputs and can be used to simulate the proxy variable of interest by proxy class. By making this forward modeling toolbox publicly available, PRYSM provides an accessible platform that maximizes the utility of proxy data and facilitates proxy-climate (simulated or historical) comparisons. Many of these codes have been employed in past studies; we review modeling approaches for each proxy class, and compare results when forced with an isotope-enabled climate simulation. Applications of multi-proxy forward modeling including parameter estimation, the effects of physical processes (such as karst transit times or firn diffusion in ice cores) on the simulated climate signal, as well as explicit modeling of time uncertainties are used to demonstrate the utility of PRYSM for a broad array of climate studies.


PRYSM | Ice Core | Ice Cor
Rod
127.0.0.1:8000/icecore/69/
You may update any of this data set's attributes:
Label:
Ice Core Test Data
Latitude:
78.00
Longitude:
1.00
Start Year:
1000
Description (Optional, 500 chars maximum):
Sample Ice Core Data
Quelccaya
This data set is public:
Public data sets are visible by anyone
Analytical Uncertainty/Measurement Error (default is ±0.1 ‰):
0.10
Age Uncertainty
Asymmetric miscounting?
Symmetric or Miscounted (default: 0.01):
0.01
Edit Attributes
You may update any individual fields by providing a .csv or .npy file:
Temperature .csv or .npy file:
Choose File
No file chosen
Accumulation .csv or .npy file:
Choose File
No file chosen
Depth .csv or .npy file:
Choose File
No file chosen
Depth Horizons .csv or .npy file:
Choose File
No file chosen
δ18OPRECIPITATION .csv or .npy file:
Choose File
No file chosen
Update Files


PRYSM | Ice Core | Ice Cor
127.0.0.1:8000/icecore/69/
Rod
PRYSM
Ice Core
BROWN
Welcome, test
Log out
Home
Cellulose
Coral
Ice core
Speleothem calcite
Data set: Ice Core Test Data
Coordinates: 78.00, 1.00
Time span: 1000 – 2005
Analytical Uncertainty: 0.10
Age Uncertainty: Symmetric – 0.01
Description: Sample Ice Core Data Quelccaya
Privacy: public
Download this data set
Plotting options:
Show time series:
Simulated Ice Core Time Series + BAM simulated Age Perturbations [2% Error Rate]
−10.6
−10.8
−11.0
−11.2
δ18OI

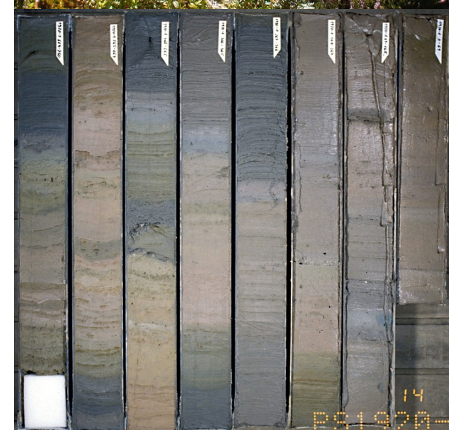

PRYSM | Coral | Test Data
Rod
127.0.0.1:8000/coral/15/
Download this data set
Plotting options:
Show time series:
Simulated Coral Time Series + BAM simulated Age Perturbations [4% Error Rate]
0.5
0.4
0.3
0.2
0.1
0.0
−0.1
−0.2
−0.3
δ18OC
1000
1200
1400
1600
1800
Year
δ18OC
1000 Age-Perturbed Realizations, [2.5 97.5 CI]
Show power spectrum:

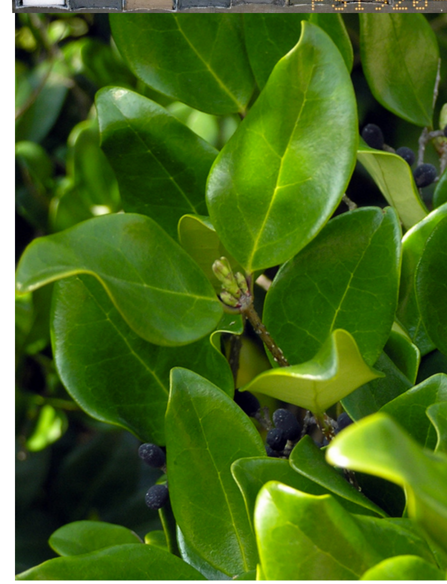

127.0.0.1:8000/coral/15/
δ18OC
1000 Age-Perturbed Realizations, [2.5 97.5 CI]
Show power spectrum:
psd.pdf
1 / 1
ENVIRONMENT
NINO3.4 SST
SSS (PSU)
PSD
Period (Years)
SENSOR
Coral δ18OC
PSD
Period (Years)
OBSERVATION
Coral δ18OC
1000 Age-Perturbed Realizations [4%], CI
PSD
Period (Years)

PRYSM v2.0

in the works

- marine and lacustrine indicators (e.g. leaf waxes) + bioturbation + compaction
  - Bronwen Konecky & Jess Tierney
- VS-Lite, translated to python
- peat?!?
- discussion:
  - what should this toolbox do in future extensions?
  - what is the most appropriate platform?
  - continuity and funding





A brief outline:
developing best practices for model-data comparison

- **Proxy System Modeling progress:**
  - Open-source, public PSM tools: PRYSM v1.0
- **Applications in Data Model Comparison**
  - Data Assimilation and Paleoclimate Reanalyses with PSMs
  - Investigating parametric uncertainties ~ checking our understanding of the proxy system
  - Data-model comparison in the frequency domain using PSMs

Applications 1: Data Assimilation and Paleoclimate Reanalyses with PSMs

Last Millennium Reanalysis Project

Applications 1: Data Assimilation and Paleoclimate Reanalyses with PSMs

Last Millennium Reanalysis Project

Journal of Geophysical Research: Atmospheres Explore this journal >

Browse Accepted Articles

Accepted, unedited articles published online and citable. The final edited and typeset version of record will appear in future.

Research Article

The Last Millennium Climate Reanalysis Project: Framework and First Results†

Gregory J. Hakim, Julien Emile-Geay, Eric J. Steig, David Noone, David M. Anderson, Robert Tardif, Nathan Steiger, Walter A. Perkins

Data Assimilation + PSMs to test common assumptions in Paleoclimate:

(1) climate proxies can be modeled as linear, univariate recorders of temperature

(2) structural errors in GCMs can be neglected.

Applications 1: Data Assimilation and Paleoclimate Reanalyses with PSMs

Kalman Gain (weights)

The data assimilation equation: $$x_a = x_b + K[y - \mathcal{H}(x_b)]$$

analysis/update prior obs/proxies PSM
"innovation"

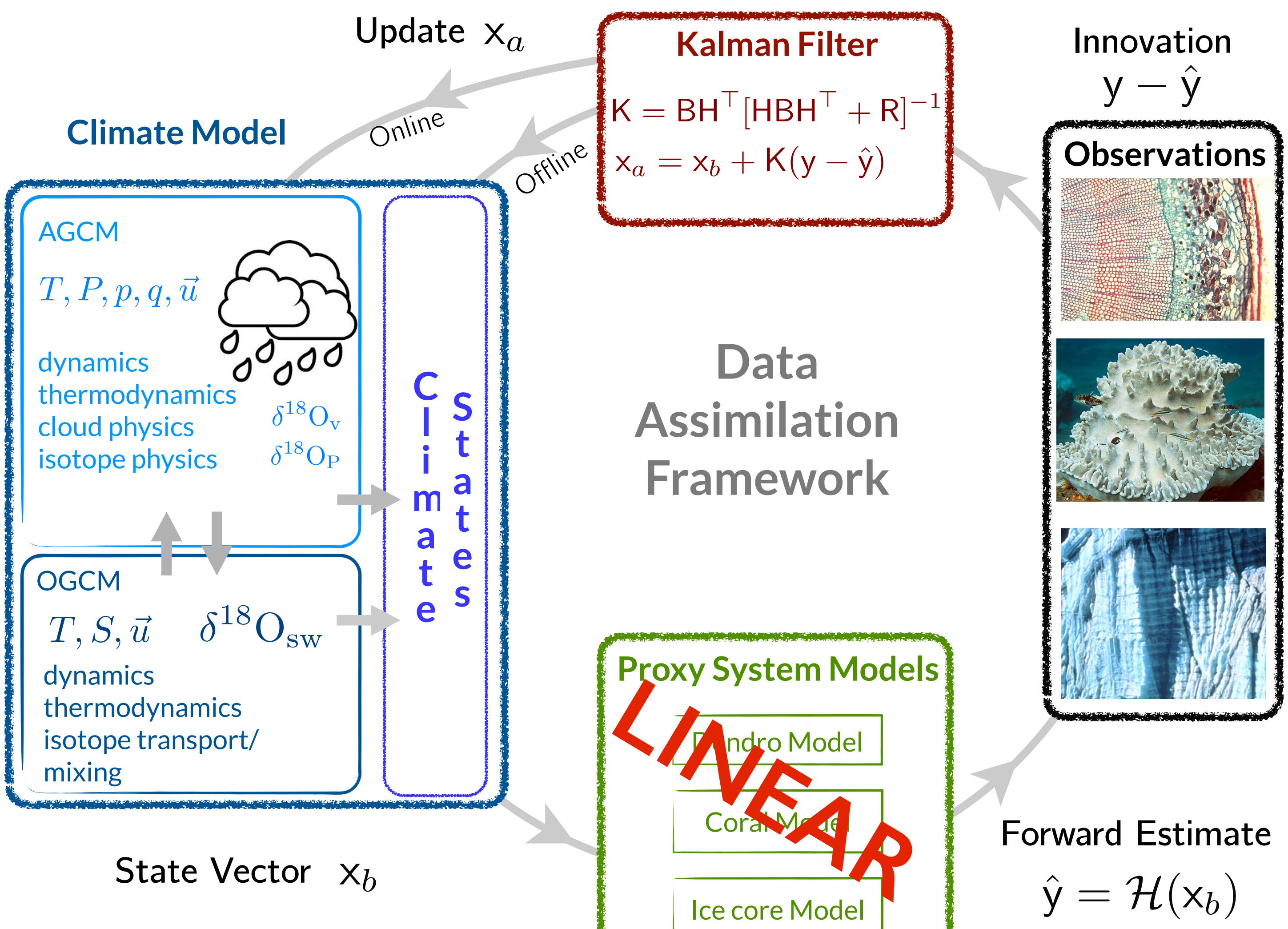
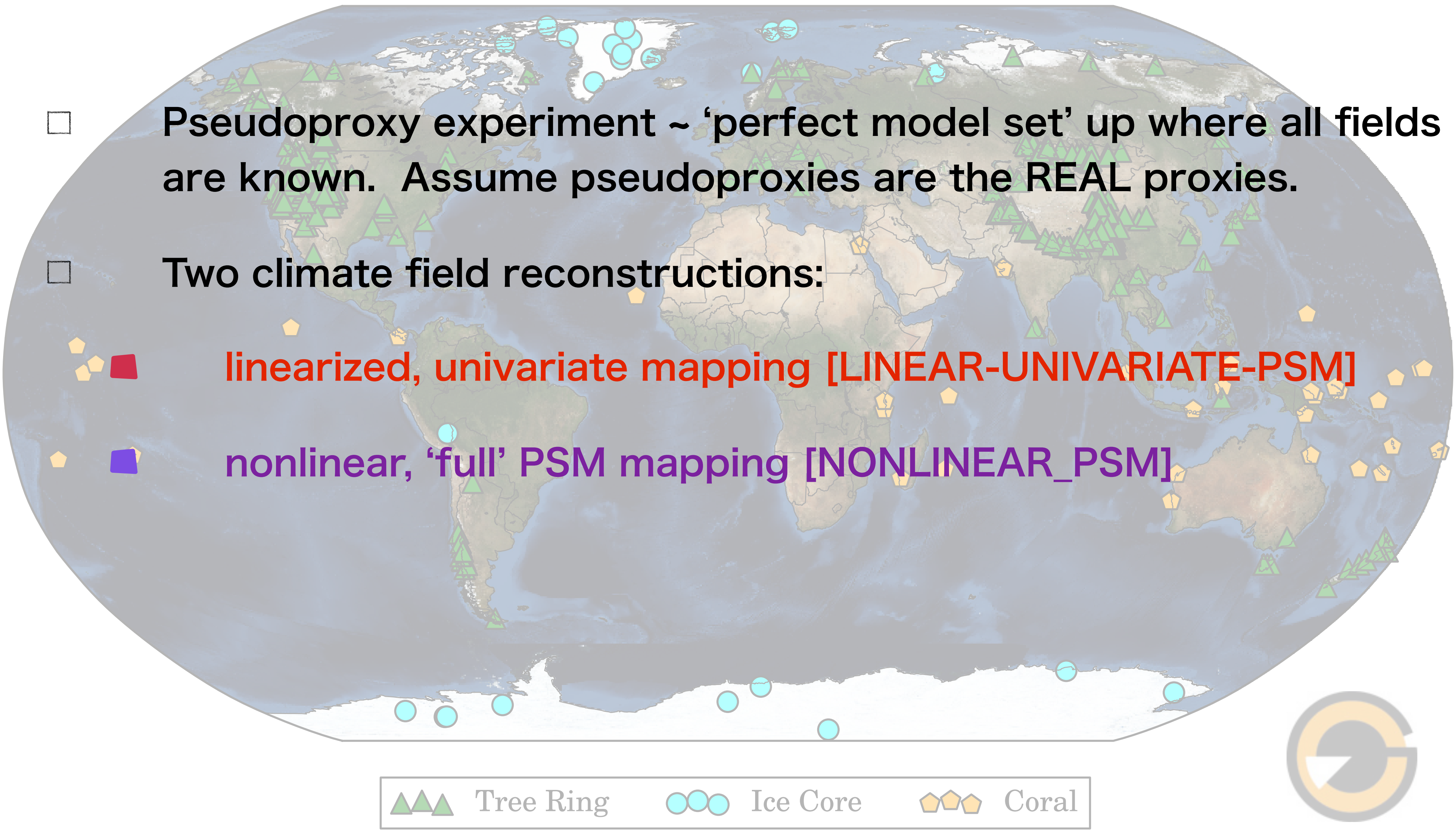


Figure Courtesy of Julien Emile-Geay

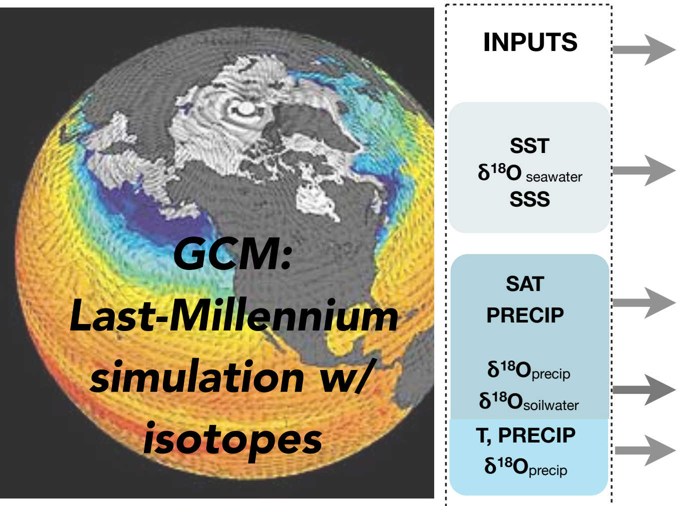
Question (1): Can climate proxies be modeled as linear, univariate recorders of temperature?

- Pseudoproxy experiment ~ ‘perfect model set’ up where all fields are known. Assume pseudoproxies are the REAL proxies.
- Two climate field reconstructions:
  - linearized, univariate mapping [LINEAR-UNIVARIATE-PSM]
  - nonlinear, ‘full’ PSM mapping [NONLINEAR_PSM]



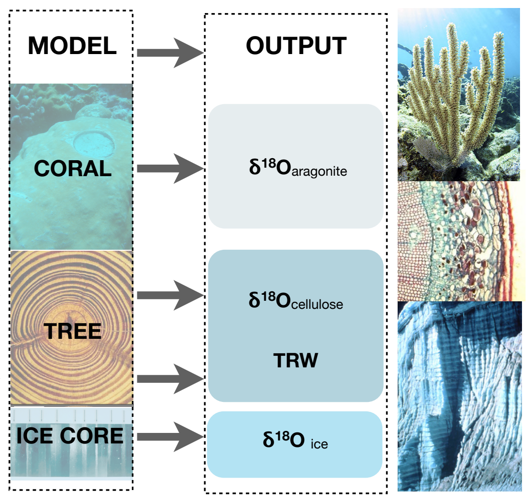
Experimental Design

1. Known Climate Fields

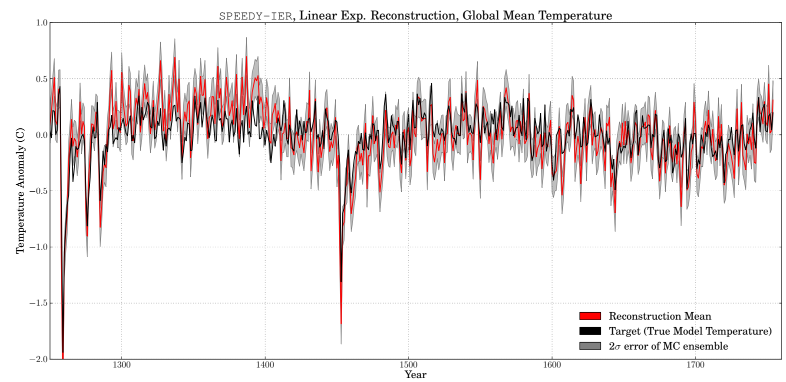


SPEEDY-IER

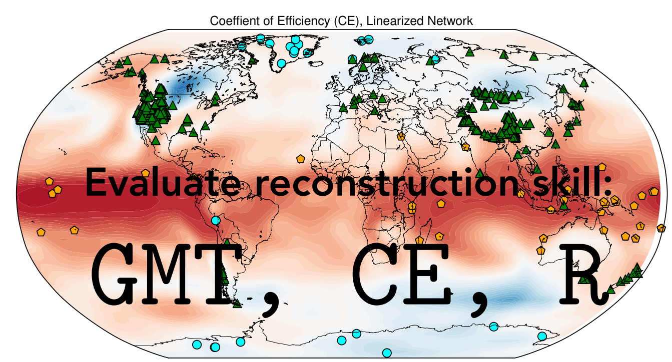
2. Generate 'Pseudo' Proxy Data



3. Reconstruct temperature *(DA)* for each year using noisy pseudo proxies



4. How good is the climate field reconstruction? Compare reconstructed temperature to original, 'true' model temperature (known, unlike in nature!)

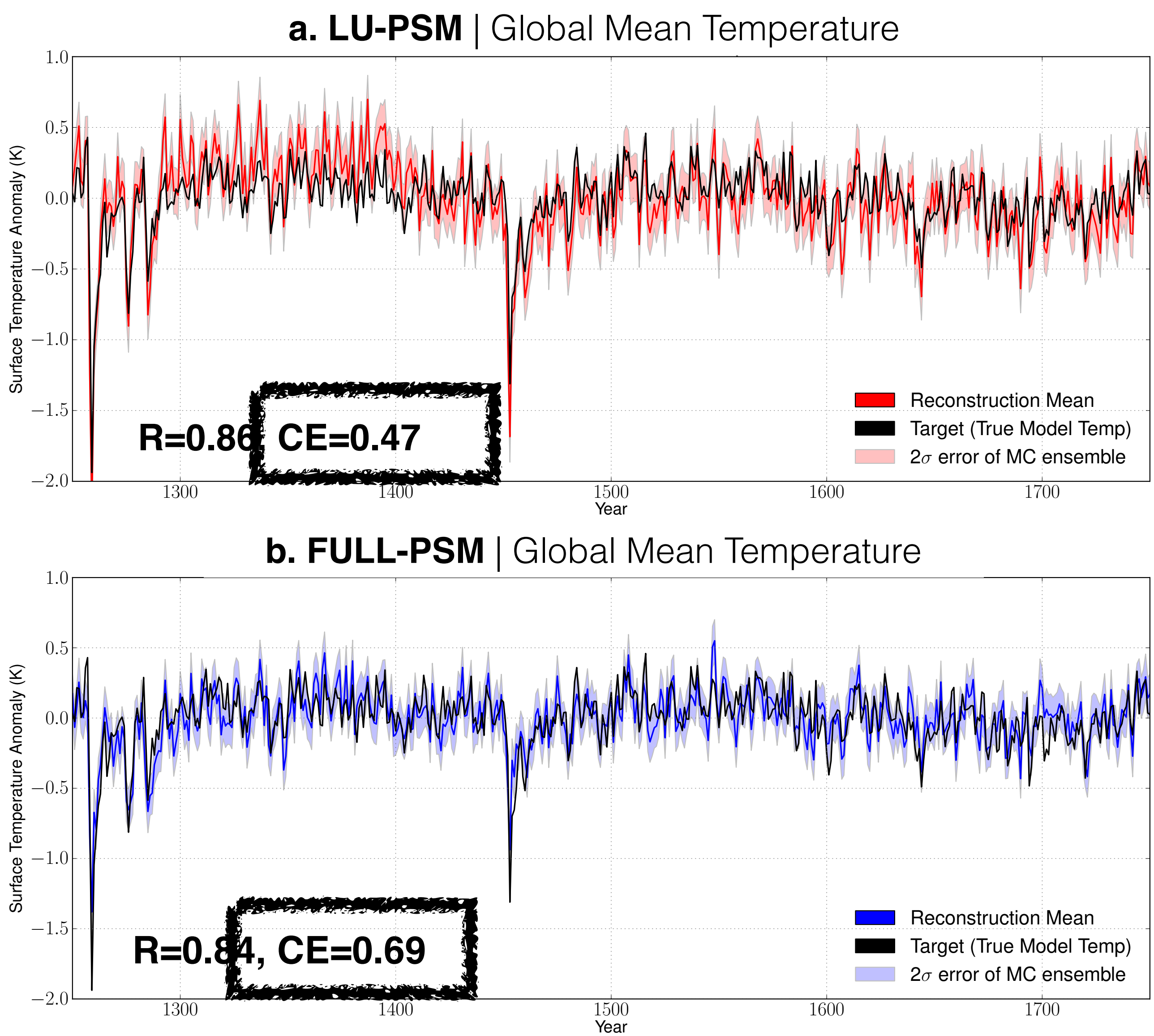


☐ *REPEAT FOR TWO PSEUDOPROXY EXPERIMENTS*

- *linearized, univariate mapping [LINEAR PSM]*
- *nonlinear, 'full' PSM mapping [NONLINEAR PSM]*

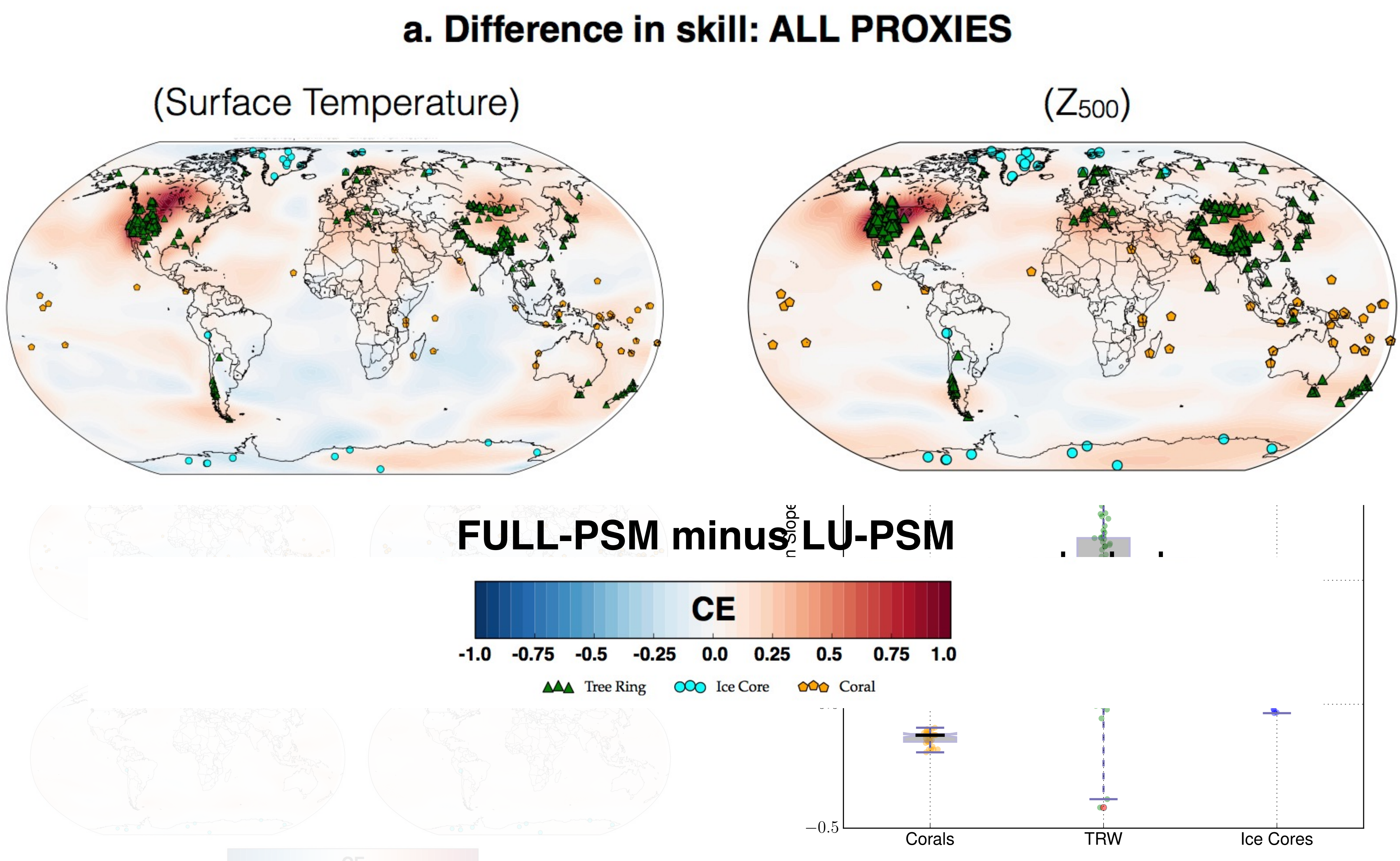
Applications 1: Data Assimilation and Paleoclimate Reanalyses with PSMs

Results 1: Linear, Univariate Models:



Dee, Steiger, Emile-Geay, Hakim, revised, JAMES

Applications 1: Data Assimilation and Paleoclimate Reanalyses with PSMs

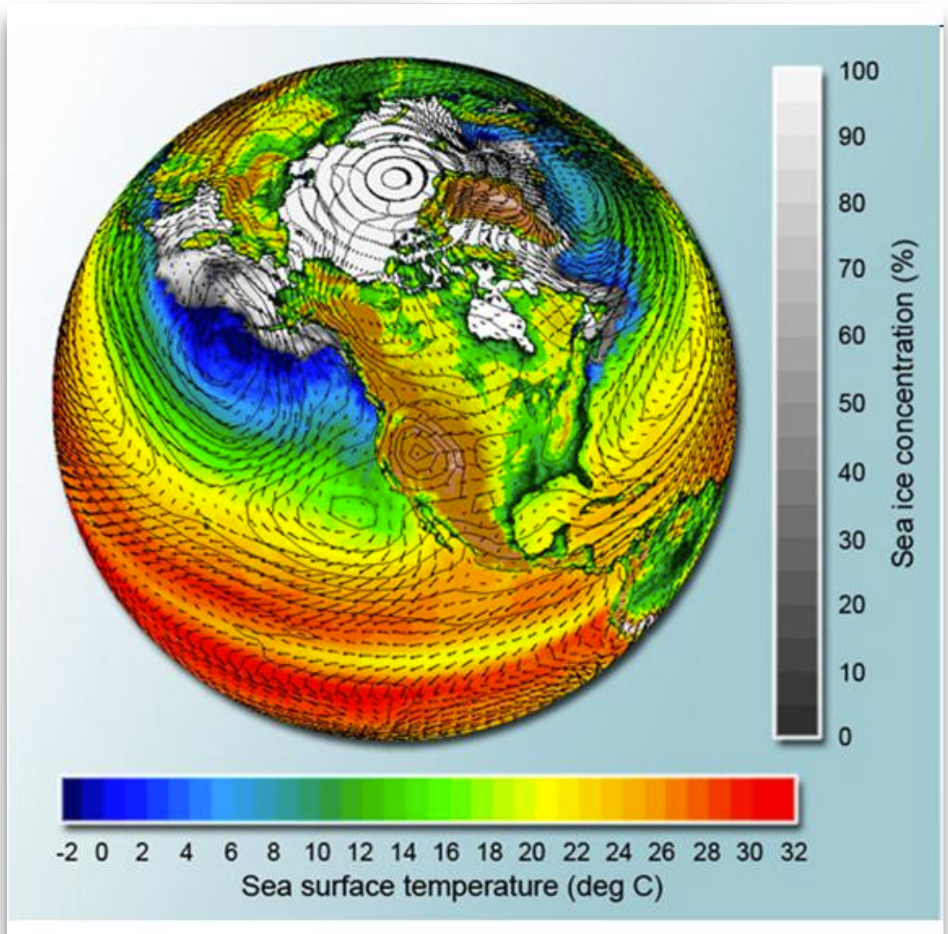


Dee, Steiger, Emile-Geay, Hakim, revised, JAMES

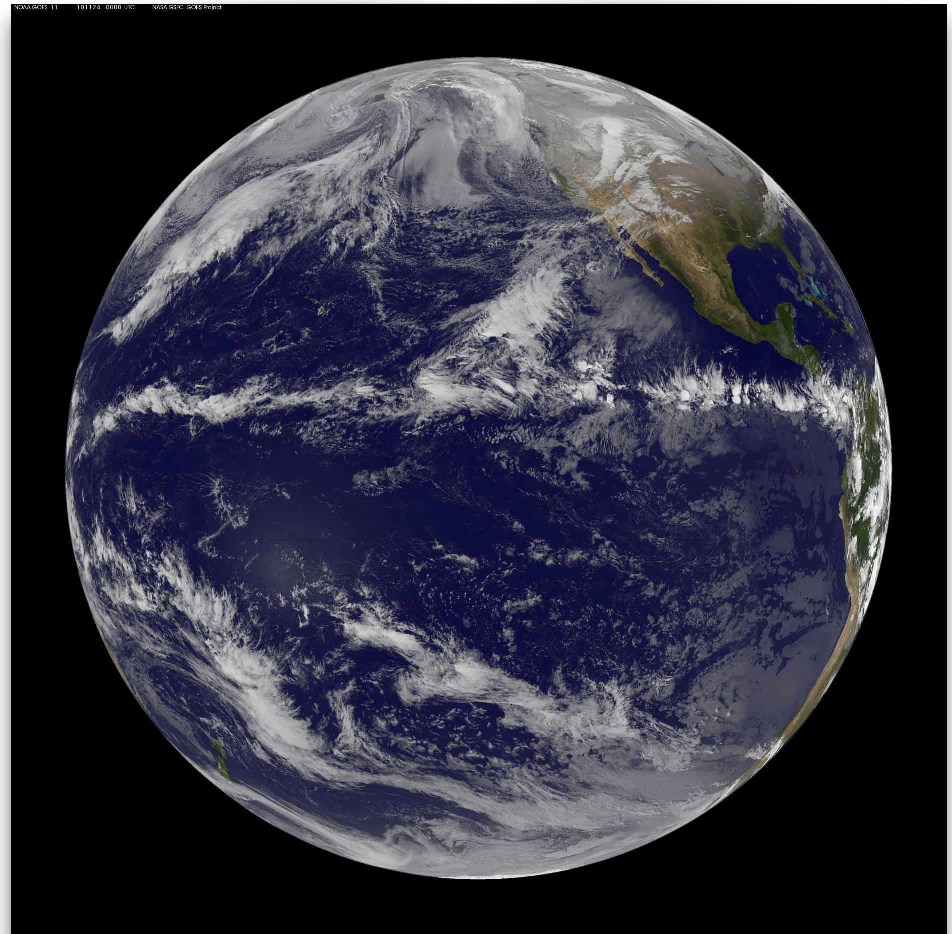
Question (2): what is the impact of GCM structural errors?

GCMs: biased

Reality: “unbiased”



≠

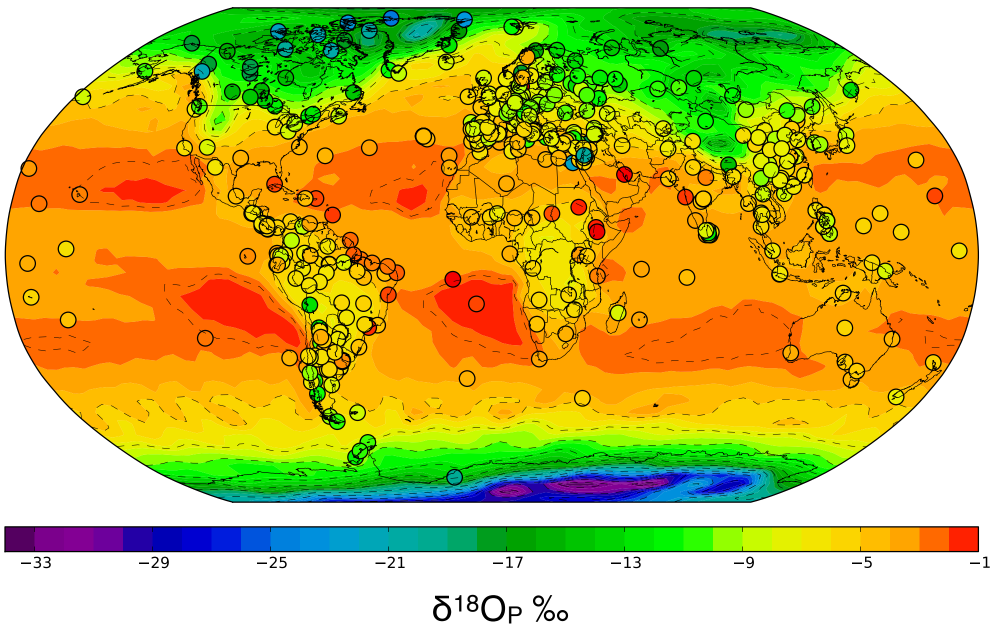


GCMs are an imperfect representation of nature, and house errors

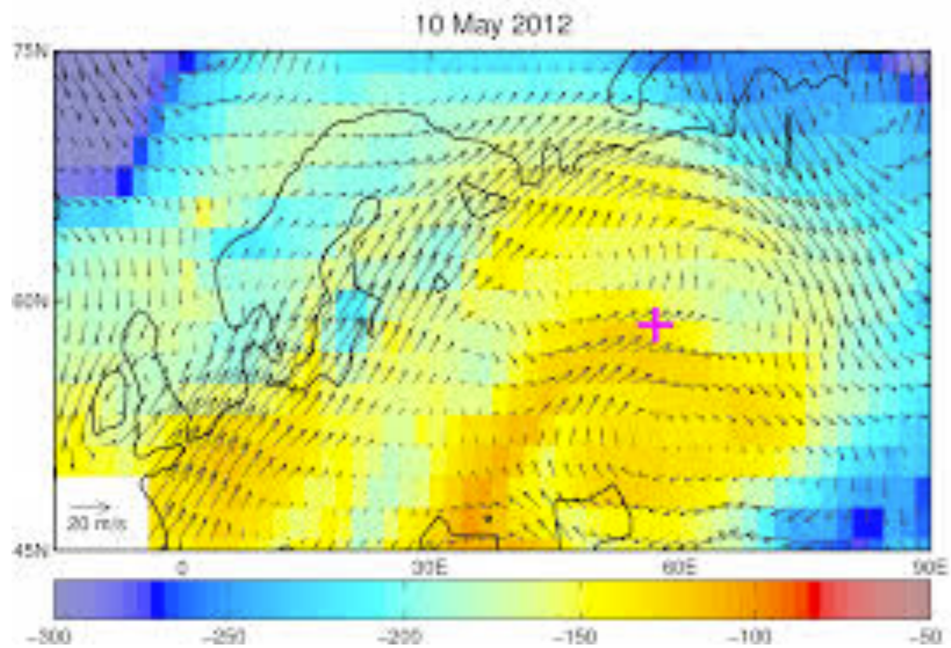
Question (2): what is the impact of GCM structural errors?

SPEEDY-IER ~ GCMs

ECHAM5-wiso ~ Reality



≠



We use ECHAM5 to approximate ‘nature,’ and try to reconstruct climate using ECHAM5-generated proxies with a SPEEDY prior.

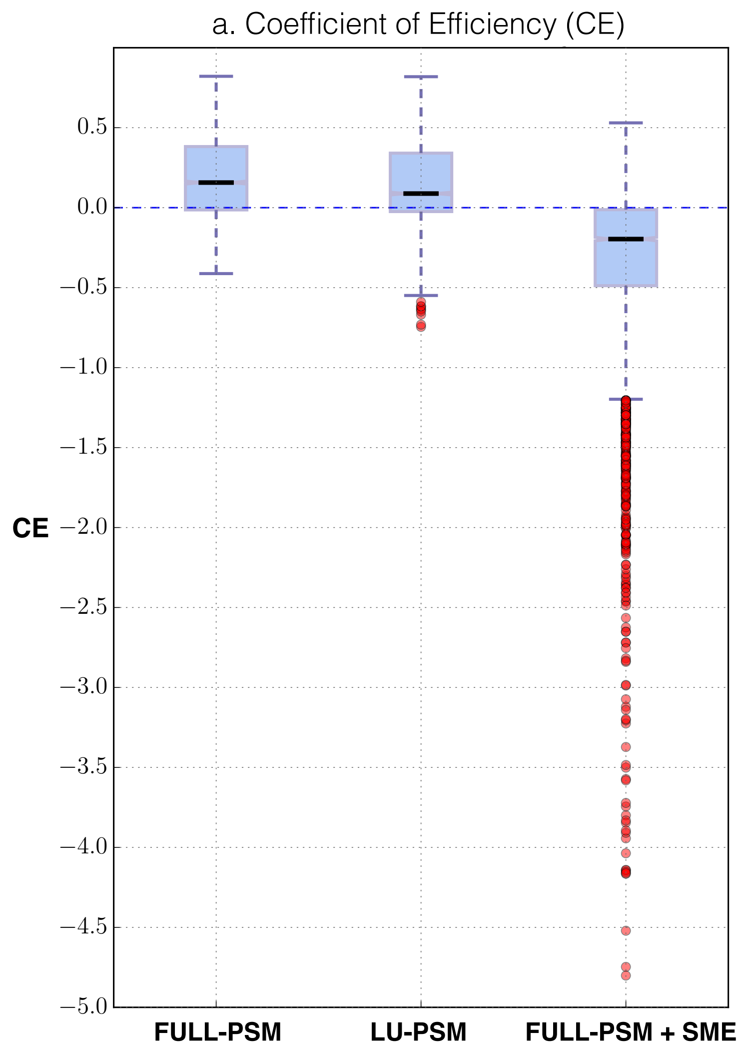
Applications 1: Data Assimilation and Paleoclimate Reanalyses with PSMs

Results 2: Structural Uncertainties in GCMs

Despite the improvement using FULL-PSM, errors in GCMs propagate forward through PSMs and may reduce reconstruction skill.

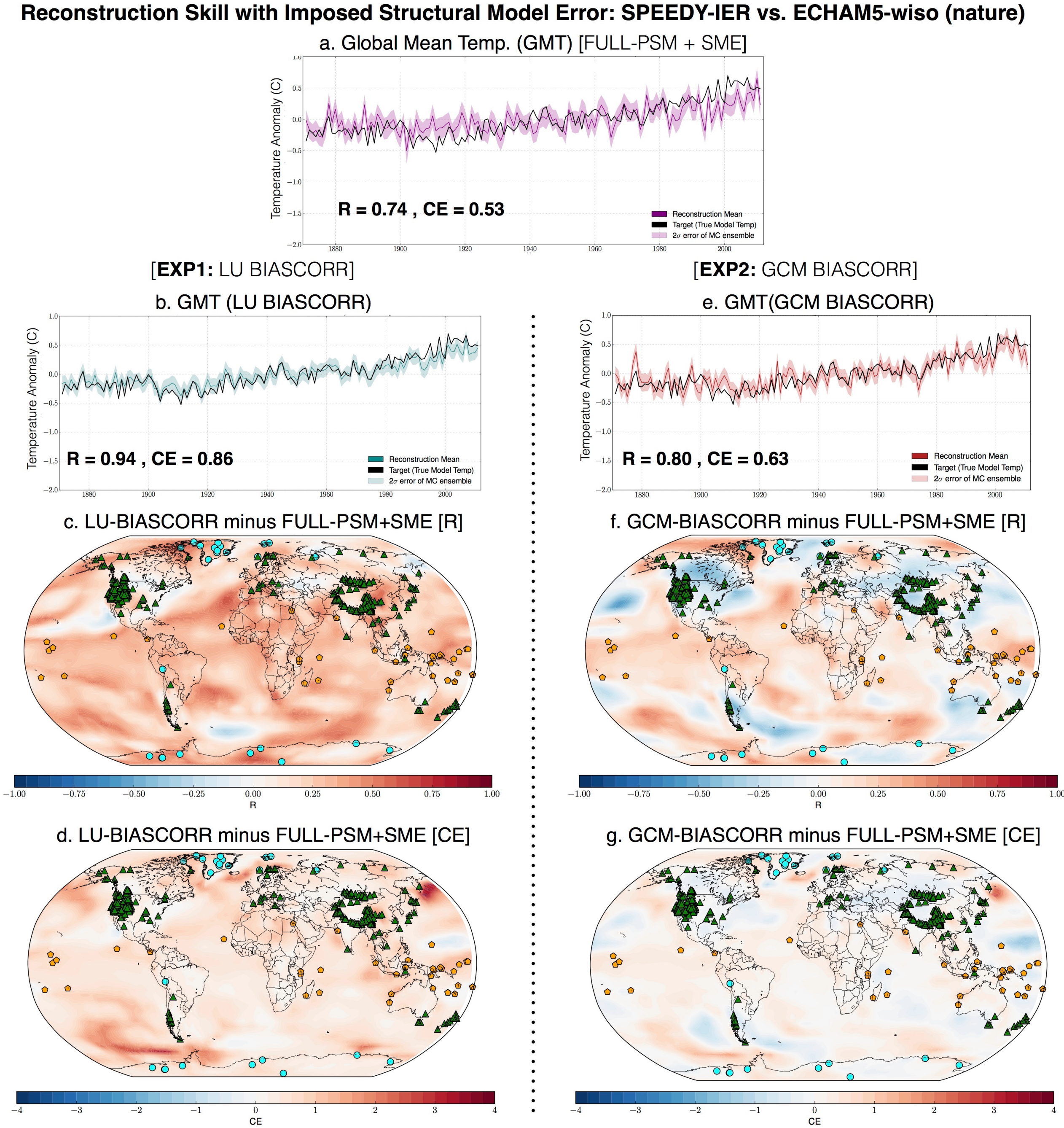
(Not true for LU-PSM, which is calibrated to the 'true' model state).

Comparison of Skill Scores:
Surface Temperature



Dee, Steiger, Emile-Geay, Hakim, revised, JAMES

Applications 1: Data Assimilation and Paleoclimate Reanalyses with PSMs



2: Structural Uncertainties

Despite the improvement using FULL-PSM, errors in GCMs propagate forward through PSMs and may reduce reconstruction skill. (Not true for LU-PSM, which is calibrated to the 'true' model state).

Two mitigating strategies:

1. Back to Linear Calibration
2. **Bias Correct the GCM**

Dee, Steiger, Emile-Geay, Hakim, revised.

Applications 1: Data Assimilation and Paleoclimate Reanalyses with PSMs

* ***The utility of PSMs for Data-Assimilation Based Reconstruction Techniques:***

  - *PSMs provide a physically-based estimate from GCMs to compare with observations.*

  - *Skill added using nonlinear models increases with increasing proxy sensitivity to variables other than temperature.*

  - *tructural uncertainties, which may prove prohibitive for using PSMs h DA and systematically reduce reconstruction skill, may be bias-correcting GCMs*

LM-ISO-MIP?

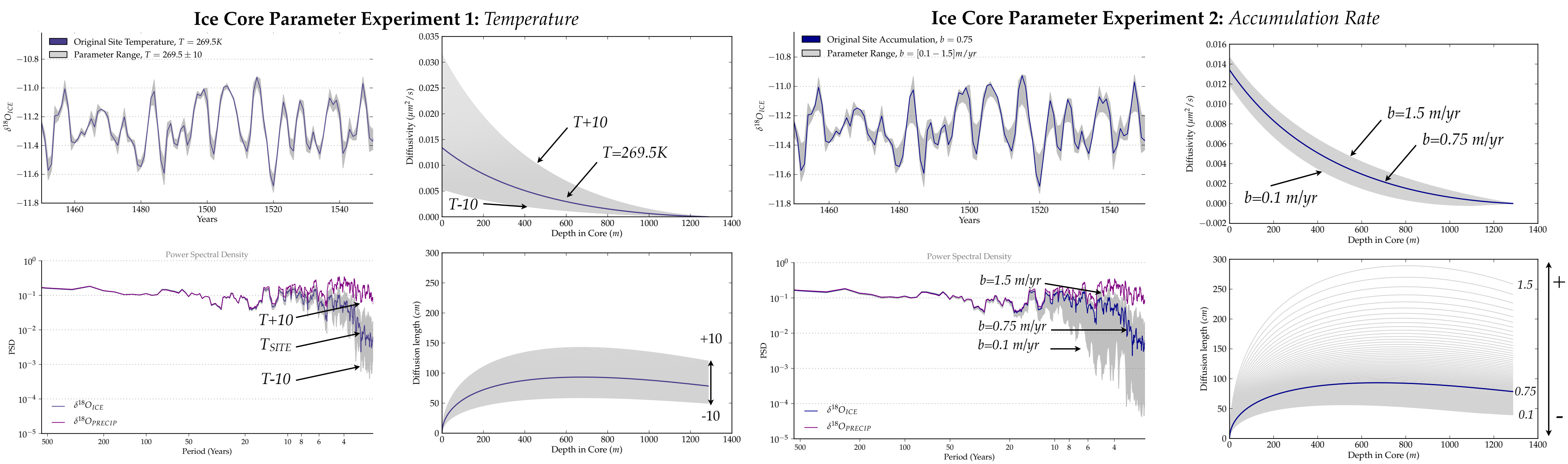
t: need to repeat analysis with multiple isotope-enabled GCMs!

Follow up questions: what is the fall out if our models for how proxies work are slightly wrong? (PSM parameter uncertainties)

Applications 2: Investigating PSM Parametric Uncertainties

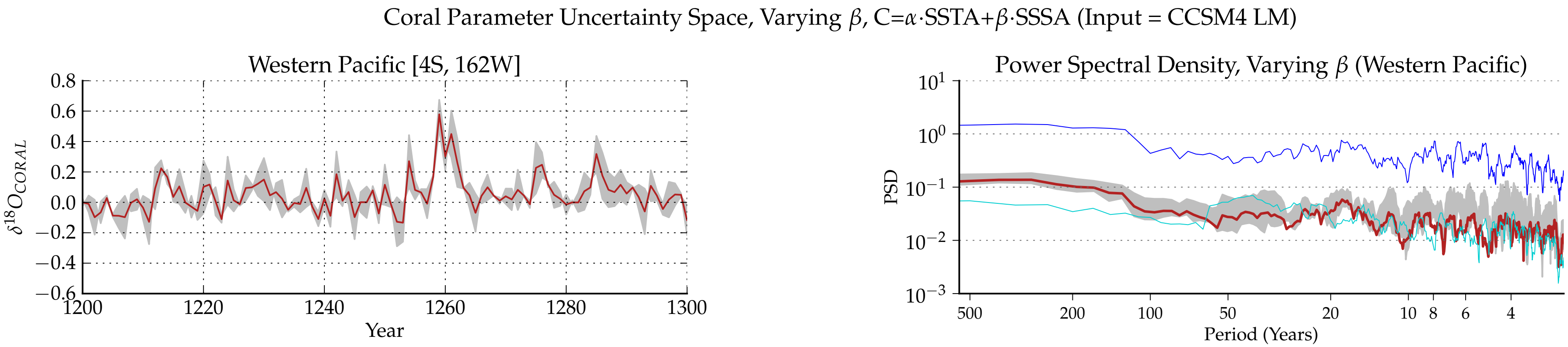
- **checking our understanding of the proxy system**

❋ Parametric uncertainties exist in our representation of proxy systems; we can use PSMs to constrain these uncertainties.



WHAT IS THE CONTRIBUTION OF EACH CLIMATE INPUT TO THE FINAL SIGNAL?

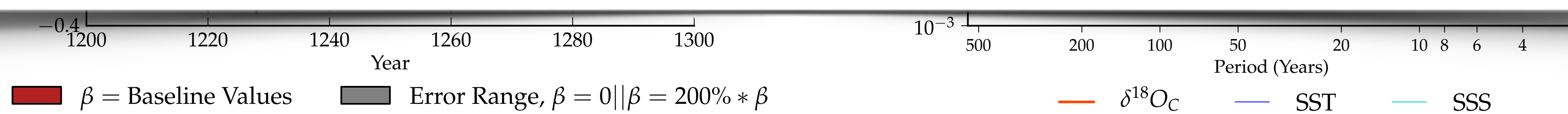
Applications 2: Investigating PSM Parametric Uncertainties



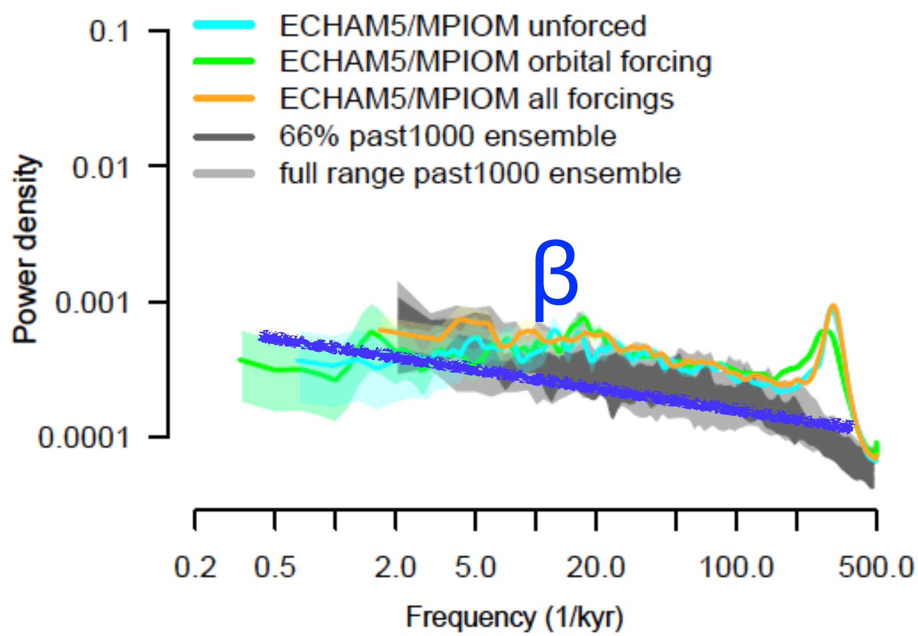
PSMs allow us to:

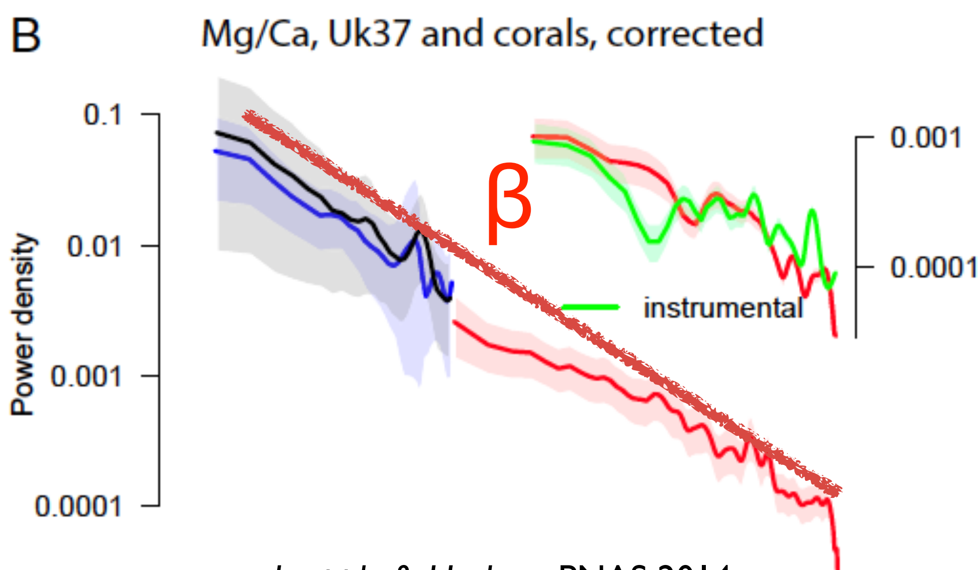
- evaluate the contribution of each input climate variable (and its variability) on the final measured signal
- quantify uncertainties in signal interpretation

Related: how do proxy system processes affect the **spectrum of variability** observed in paleoclimate data?



Applications 3: Enhanced Data-Model Comparison with PSMs





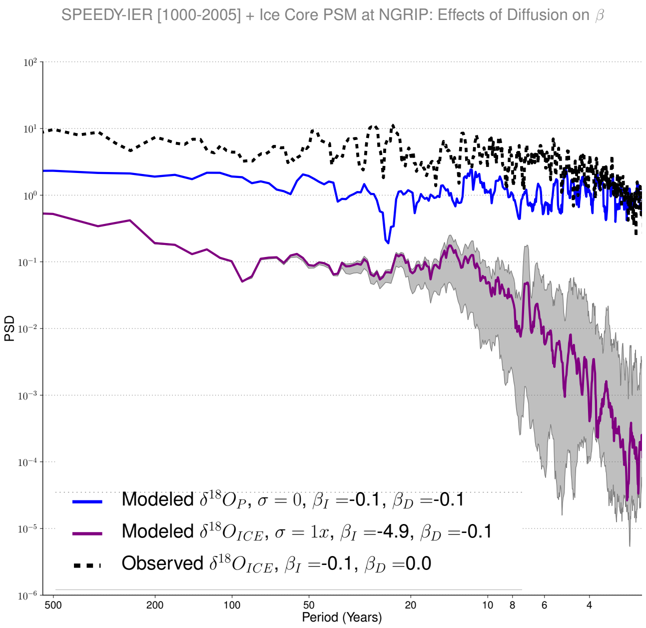
Laepple & Huybers, PNAS 2014

- Data-model comparison in the frequency domain (building on Thomas Laepple, Toby Ault's work, but from the forward direction). [see Laepple & Huybers, PNAS 2014, Ault et al., 2013]

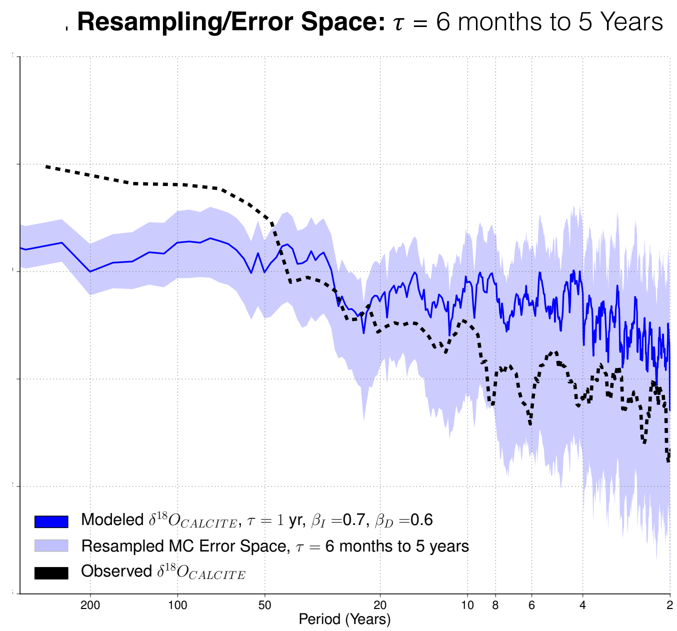
How does this comparison change when GCM simulated climate fields are converted to proxy units using PSMs?

Applications 3: Enhanced Data-Model Comparison with PSMs

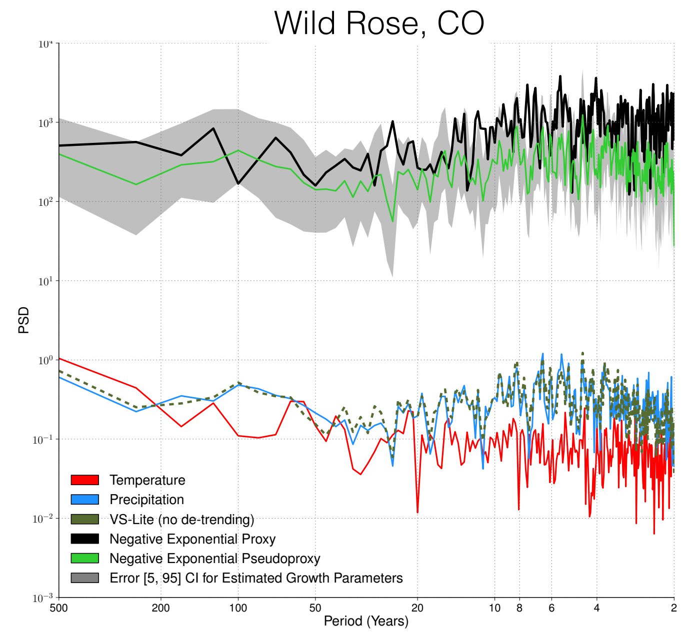
What processes *inherent to the proxy system itself* can alter its power spectrum?



Ice Cores: Diffusion and Compaction



Speleothems: Groundwater Residence Time



TRW: soil moisture and detrending method

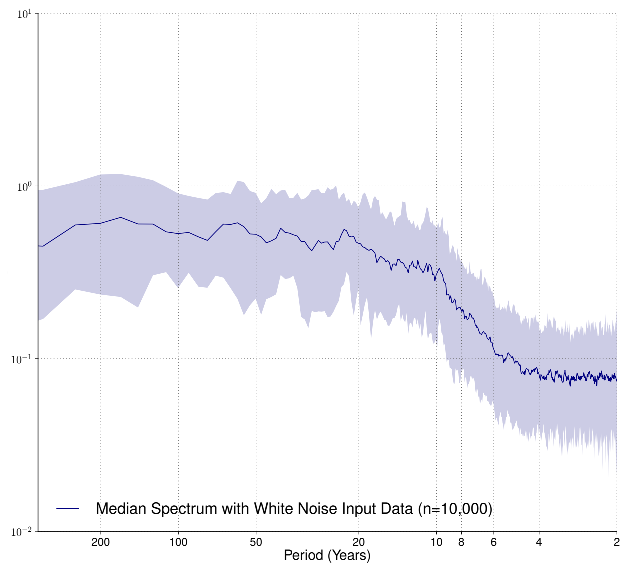
Processes such as diffusion, karst water storage, soil moisture seasonality/ memory and detrending all have an impact on the proxies' power spectra.

Dee et al., in prep

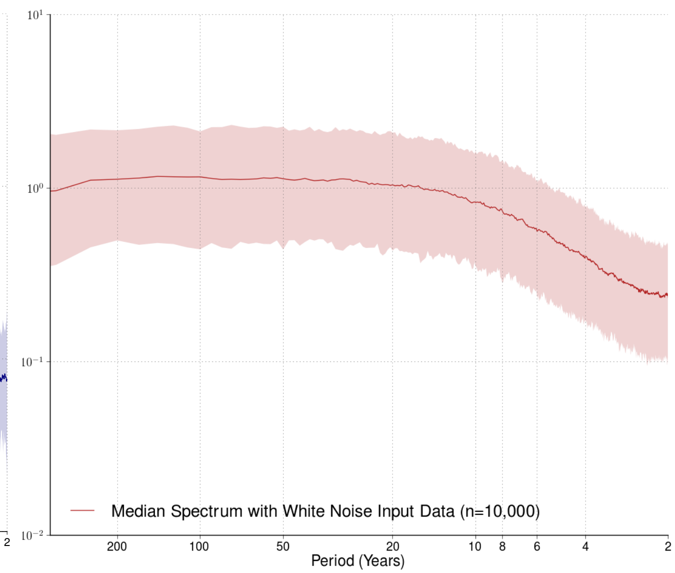
Applications 3: Enhanced Data-Model Comparison with PSMs

“Spectral Fingerprints” By Proxy Type

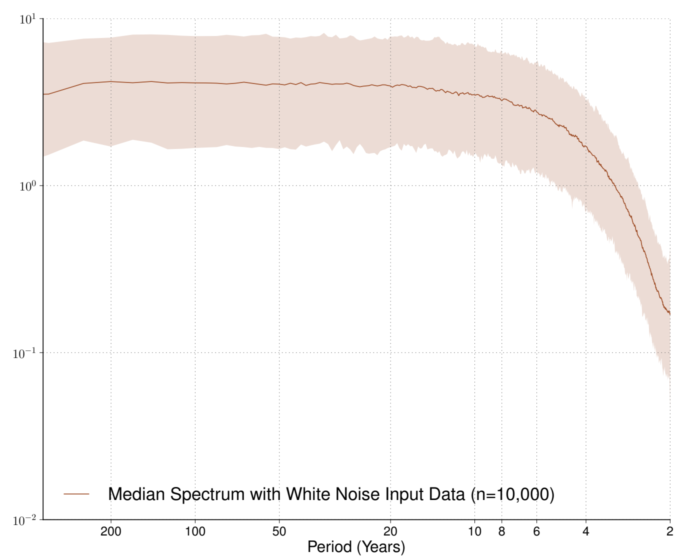
a. Ice Core $\delta^{18}O$: $\beta_I = 0.86$, $\beta_D=0.26$



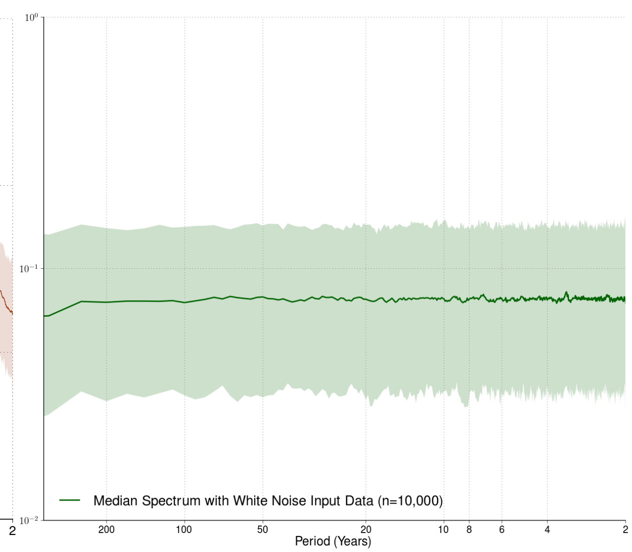
b. Speleothem $\delta^{18}O_{CALCITE}$: $\beta_I = 0.91$, $\beta_D=0.12$



c. Tree Ring Width: $\beta_I = 1.53$, $\beta_D=0.06$



d. Tree Ring $\delta^{18}O_{CELLULOSE}$: $\beta_I =-0.005$, $\beta_D=-0.004$

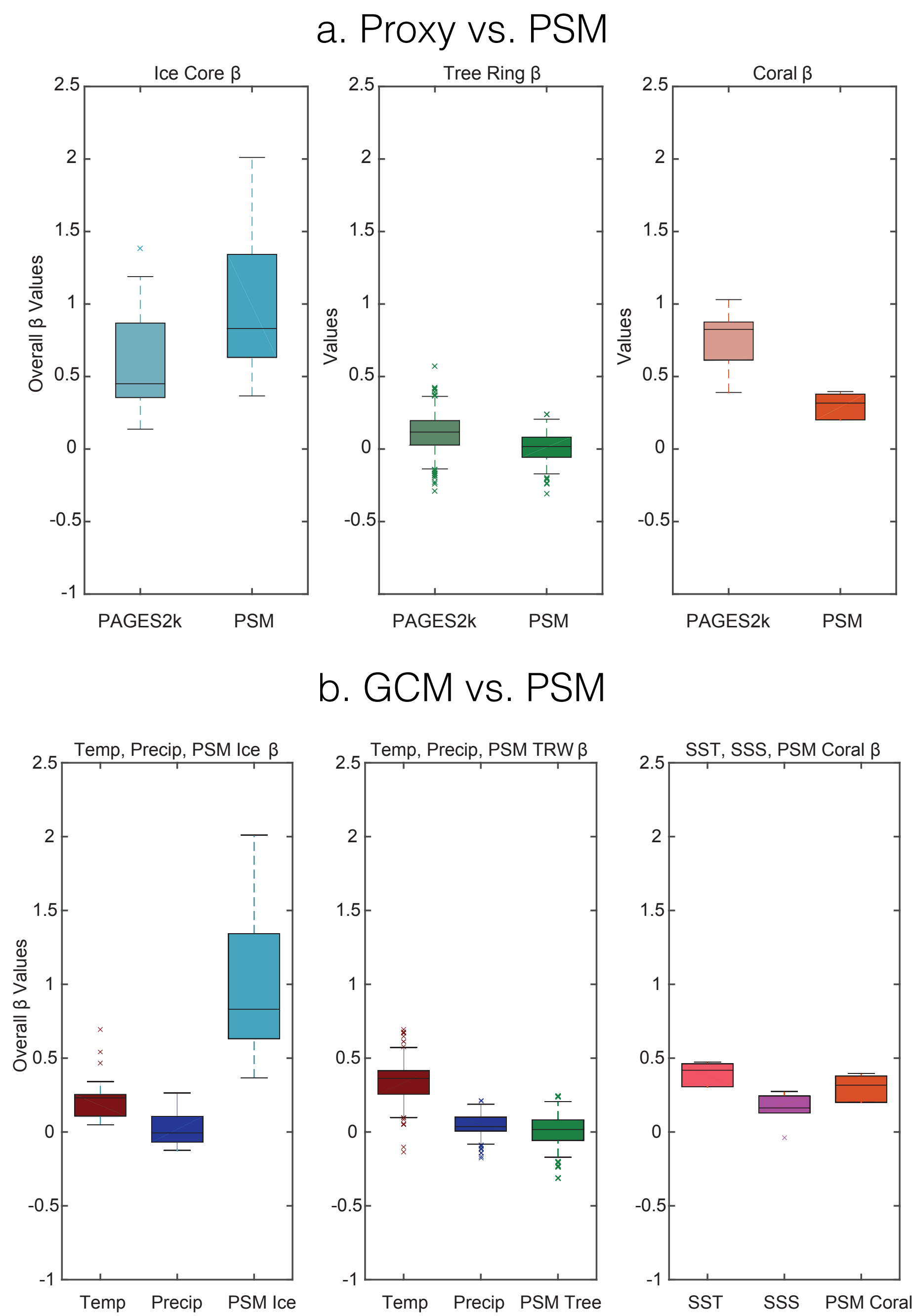


Dee et al., in prep

Applications 3: Enhanced Data-Model Comparison with PSMs

Bottom Line:

converting climate model output to proxy units helps us compare models and data in a more meaningful way.



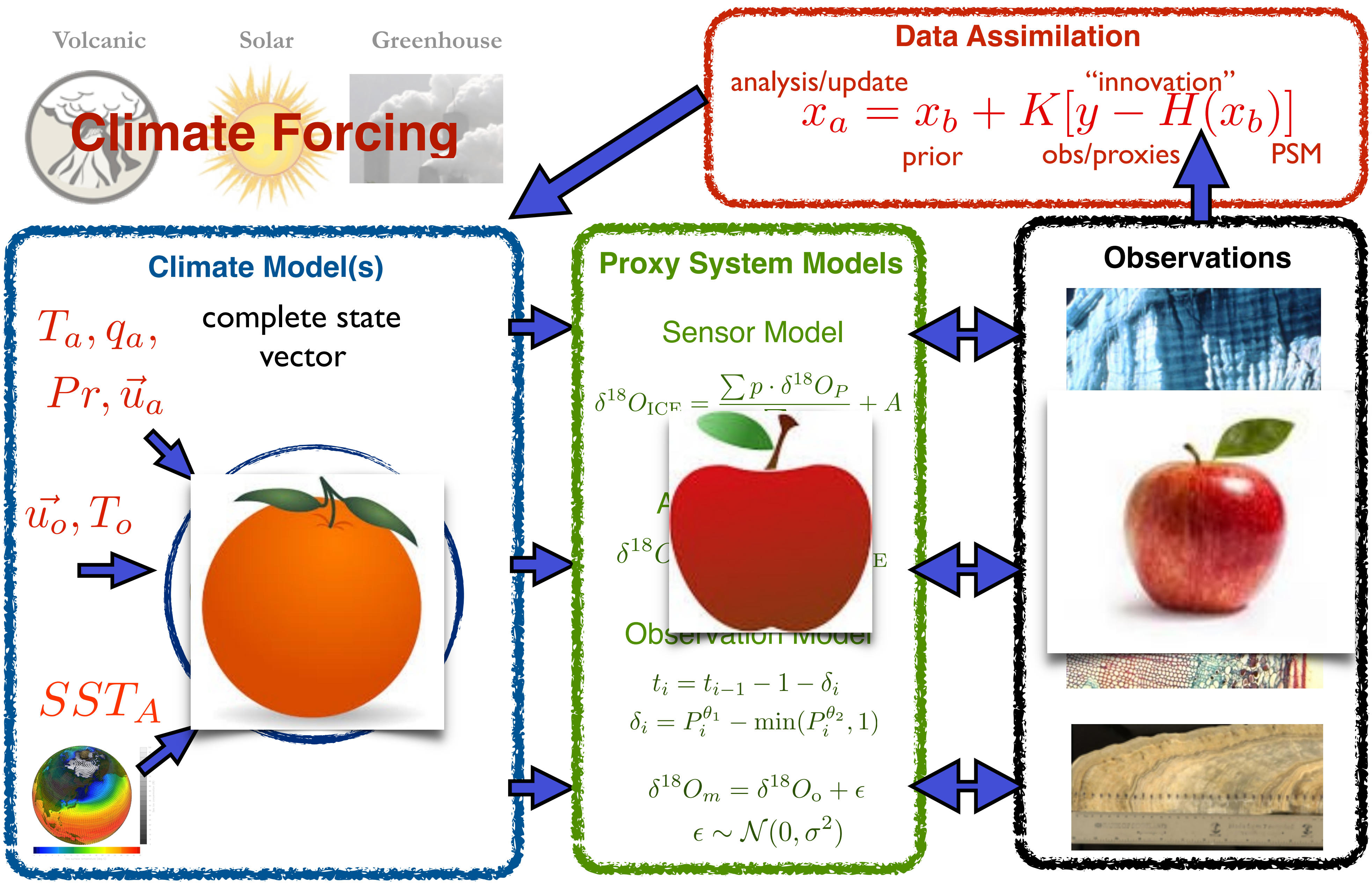
Dee et al., in prep

Discussion Points: **Challenges** in Proxy System Modeling

- Complexity and design choices:
  - How ‘fancy’ does the model have to be do be realistic and yield valuable insight?
  - How do these fanciness choices differ across:
    - proxy types
    - location / regional vs. global scale (tropical vs. mid-lat vs. high lat)
    - resolution in time and space
- Compatibility with instrumental data and GCM output (PSM inputs)
- **>> PSM session at AGU** (w/Bronwen Konecky and Corinne Wong): “Advances in proxy system modeling and data-model comparison” (PP)

A Formal Data-Model Comparison Strategy

Forward Climate-to-Proxy Modeling GCM + PSM



Thank you!

- Coauthors ~ **Data Assimilation w/PSMs:** Nathan Steiger, Julien Emile-Geay, Greg Hakim
- Coauthors ~ **Data-Model Comparison with PSMs:** Luke Parsons, Garrison Loope, Toby Ault, Jonathan Overpeck

Future Work: PSM Applications

- Improved paleoclimate signal interpretation
- Tracking external forcing from climate to proxy
- Sensor Placement
- Categorizing extremes in hydroclimate variability

How does this proxy work?

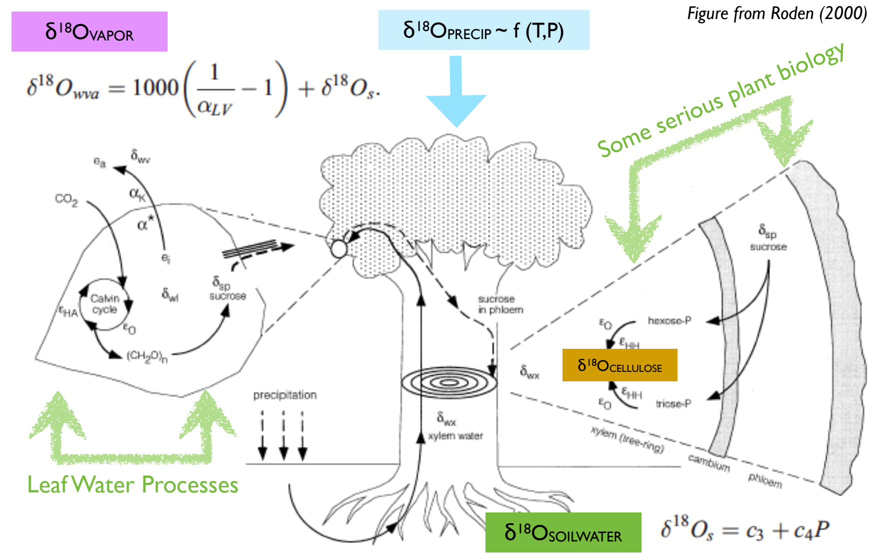


Fig. 1. A diagram of the isotopic fractionation events occurring between precipitation input and tree-ring cellulose.

[Evans (2007), Evans (2004), Barbour et al. (2004), Roden (2000)]